现代数学基础丛书 177

均匀试验设计的理论和应用

方开泰　刘民千　覃　红　周永道　著



现代数学基础丛书 177

均匀试验设计的理论和应用

方开泰　刘民千　覃　红　周永道　著

科 学 出 版 社

北　京

内 容 简 介

本书介绍均匀试验设计的理论、方法和应用. 均匀设计是一类模型未知的部分因子设计、计算机试验中的空间填充设计、超饱和设计或存在模型误差的稳健设计, 该方法也可以应用于混料试验. 本书包括在不同试验区域上的均匀性度量、构造均匀设计的确定性和随机性方法、均匀设计相应的建模方法、均匀性与其他设计准则之间的联系及均匀性的应用及混料均匀设计等内容.

本书可作为高等院校统计学专业研究生的教材, 也可供实验工作者、相关专业的研究生和教师参考, 还可作为社会科学、问卷调查设计人员的参考书.

图书在版编目(CIP)数据

均匀试验设计的理论和应用/方开泰等著. —北京：科学出版社，2019.10
(现代数学基础丛书; 177)

ISBN 978-7-03-062444-4

Ⅰ.①均…　Ⅱ.①方…　Ⅲ.①均匀性-试验设计　Ⅳ.O212.6

中国版本图书馆 CIP 数据核字(2019) 第 215326 号

责任编辑: 李静科　李香叶 / 责任校对: 邹慧卿
责任印制: 吴兆东 / 封面设计: 陈　敬

科学出版社出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京厚诚则铭印刷科技有限公司 印刷

科学出版社发行　各地新华书店经销

*

2019 年 10 月第　一　版　开本: 720 × 1000　1/16
2020 年 2 月第二次印刷　印张: 18 1/4　插页: 2
字数: 345 000

定价: 128.00 元

(如有印装质量问题, 我社负责调换)

《现代数学基础丛书》序

对于数学研究与培养青年数学人才而言, 书籍与期刊起着特殊重要的作用. 许多成就卓越的数学家在青年时代都曾钻研或参考过一些优秀书籍, 从中汲取营养, 获得教益.

20 世纪 70 年代后期, 我国的数学研究与数学书刊的出版由于“文化大革命”的浩劫已经破坏与中断了 10 余年, 而在这期间国际上数学研究却在迅猛地发展着. 1978 年以后, 我国青年学子重新获得了学习、钻研与深造的机会. 当时他们的参考书籍大多还是 50 年代甚至更早期的著述. 据此, 科学出版社陆续推出了多套数学丛书, 其中《纯粹数学与应用数学专著》丛书与《现代数学基础丛书》更为突出, 前者出版约 40 卷, 后者则逾 80 卷. 它们质量甚高, 影响颇大, 对我国数学研究、交流与人才培养发挥了显著效用.

《现代数学基础丛书》的宗旨是面向大学数学专业的高年级学生、研究生以及青年学者, 针对一些重要的数学领域与研究方向, 作较系统的介绍. 既注意该领域的基础知识, 又反映其新发展, 力求深入浅出, 简明扼要, 注重创新.

近年来, 数学在各门科学、高新技术、经济、管理等方面取得了更加广泛与深入的应用, 还形成了一些交叉学科. 我们希望这套丛书的内容由基础数学拓展到应用数学、计算数学以及数学交叉学科的各个领域.

这套丛书得到了许多数学家长期的大力支持, 编辑人员也为其付出了艰辛的劳动. 它获得了广大读者的喜爱. 我们诚挚地希望大家更加关心与支持它的发展, 使它越办越好, 为我国数学研究与教育水平的进一步提高做出贡献.

杨 乐

2003 年 8 月

序

在科学试验与工农业生产中, 经常要做试验. 如何安排试验, 使试验次数尽量少, 而又能达到好的试验效果呢? 这是经常会碰到的问题. 解决这个问题有一门专门的学问, 叫做试验设计. 试验设计得好, 会事半功倍, 反之就有可能事倍功半了.

20 世纪 60 年代, 华罗庚教授在我国倡导与普及的 "优选法", 即国外的斐波那契方法, 与我国的数理统计学者在工业部门中普及的 "正交设计法" 是两种常用的试验设计方法. 这些方法经普及后, 已被广大技术人员与科学工作者掌握, 并取得一系列成就, 产生了巨大的社会效益和经济效益. 随着科学技术的深入发展, 上述两种方法就显得不够了. "优选法" 是单变量情形下的最佳方法, 其假定我们处理的实际问题中只有一个因子起作用. 然而, 在解决实际问题时, 这种情形几乎是没有的. 所以在使用时, 只能抓 "主要矛盾", 即突出一个因子, 而将其他因子固定, 这样来安排试验. 因此 "优选法" 还不是一个很精确的近似方法. "正交设计" 的基础是拉丁方理论与群论, 可以用来安排多因子的试验, 而且试验次数对各因子的各水平的所有组合数来说是大大地减少了, 但对于某些工业试验与昂贵的科学试验来说, 仍然因为试验次数太多而无法安排.

1978 年, 中国七机部由于导弹设计的要求, 提出了一个五因子的试验, 希望每个因子的水平数要多于 18 个, 而试验总数又不超过 50, 显然优选法和正交设计都不能用, 方开泰教授在 20 世纪 70 年代中期, 曾为近似计算一个多重积分问题找过我, 我向他介绍了多重数值积分的方法并取得了好结果, 这就使他想到是否可以将数论方法用于试验设计的问题, 于是我们经过几个月的共同研究, 提出了一种新的试验设计, 即均匀设计, 将这一方法成功地用于导弹设计. 我们的文章在 80 年代初发表后, 40 年来, 均匀设计已在我国得到较广泛的普及与使用, 取得了一系列可喜的成绩.

均匀设计属于近 60 年来发展起来的"伪蒙特卡罗方法"的范畴. 将经典的单变量问题的计算方法推广后用于多变量问题的计算时, 计算量往往跟变量个数有关. 即使计算机运算能力进步迅速, 这种方法仍无法实际应用. 乌拉姆 (S. Ulam) 与冯 • 诺依曼 (J. Von Neumann) 在 20 世纪 40 年代提出蒙特卡罗方法, 即统计模拟方法, 该方法的大意是将一个分析问题化为一个有同样答案的概率问题, 然后用统计模拟的方法来处理后者, 这样使一些困难的分析问题得以解决, 例如, 多重定积分的近似计算等. 蒙特卡罗方法的关键是找一组随机数作为统计模拟之用, 所以这一方法的精度在于随机数的均匀性与独立性.

20 世纪 50 年代末, 有些数学家试图用确定性方法寻找空间中均匀散布的点集来代替蒙特卡罗方法中的随机数, 已经找到的点集都是用数论方法找到的. 按照外尔 (H. Weyl) 定义的测度来度量, 它们的均匀性很好, 但独立性差些, 用这些点集来代替蒙特卡罗方法中的随机数, 往往会得到更精确的结果. 这一方法称为伪蒙特卡罗方法或数论方法, 数学家将这一方法成功地用于多重积分的近似计算.

从统计学的观点看, 伪随机数就是一个在立方体内均匀分布的样本. 数值积分需要大样本, 均匀设计则要找小样本. 由于这个样本比正交设计所对应的样本要均匀, 所以用它来安排试验会得到好的效果. 当然在寻求小样本时, 寻求大样本的方法是起了借鉴作用的.

均匀设计只是数论方法的一个应用, 数论方法还有广泛应用的园地. 例如, 多重插值公式的建立、某些积分与微分方程组的近似求解、求函数全局极值、求某些多元分布的近似代表点及用于统计推断的一些问题, 如多元正态性检验及多元球性检验等.

早在 50 年代末, 外国刚开始研究伪蒙特卡罗方法时, 华罗庚教授就倡议并领导了这一方法在我国的研究, 他的开拓性成果总结在我们的专著《数论在近似分析中的应用》(科学出版社, 1978 年; 英文版: Springer-Verlag and Science Press, 1981 年) 中, 这些工作是方开泰教授与我合作工作的重要背景与参考材料之一.

我与方开泰教授合作了近 40 年, 由于他既是一位数学家, 也是一位统计学家, 又有长期在中国各工业部门普及应用数理统计的宝贵经验, 所以他有很好的应用数学背景与洞察力. 他能及时地提出有价值的研究问题及解决问题的可能途径, 我们的合作既是愉快的, 又是富于成效的, 我们的成果总结在专著*Number-Theoretic Methods in Statistics* (Chapman and Hall, 1994) 之中.

方开泰教授和他的合著者的这本书着重于均匀设计的理论和应用, 也包括了他们近 20 年来的最新研究成果. 我相信该书的出版对于进一步推动均匀设计的研究将是很重要的, 我愿借此机会预祝该书成功.

王 元
中国科学院院士
2019 年 1 月

前　　言

本书的目的是介绍均匀试验设计的理论、方法和应用. 我们可以认为均匀设计是一类模型未知的部分因子设计、计算机试验中的空间填充设计、超饱和设计或存在模型误差的稳健设计. 该设计方法也可以应用于混料试验. 本书为对均匀设计理论感兴趣的读者提供均匀设计的相应理论及最新研究成果.

在多因子试验中, 试验设计非常有用且在工业生产、高科技、科学和其他领域中越显重要. 试验设计是统计学的一个重要分支, 且有悠久的历史, 其包括许多的方法及大量的设计. 关于不同类型设计的全面介绍可参见 S. Ghosh 和 C.R. Rao 编辑的*Handbook of Statistics* 第 13 卷.

经典的试验设计方法, 如部分因子设计和最优设计, 都有相应的统计模型. 部分因子设计的模型为估计因子的主效应和一些因子之间的交互效应. 最优设计考虑一个含有未知参数的回归模型. 然而, 在很多情形下, 试验人员并不知道其真实回归模型, 此时如何选择试验点是一个有挑战性的问题. 一个自然的想法是把试验点均匀地散布在试验区域. 称这样的设计方法为均匀试验设计, 或简称为均匀设计, 其由方开泰教授和王元院士在 1980 年首次提出, 参见 (Fang, 1980; Wang and Fang, 1981). 均匀设计在工业界已有数千例成功的应用案例.

许多物理现象可由线性、非线性、常微分或偏微分方程刻画, 然而其输入和输出往往没有解析表达式. 计算机试验是指通过一些软件在计算机上重现这些物理现象的方法, 是近几十年来的热门研究课题, 包括设计和建模两部分. 空间填充设计在计算机试验中起关键作用. 实际上, 均匀设计是一种空间填充设计. 本书侧重于均匀设计的有关理论, 并给出均匀设计、正交设计、组合设计、超饱和设计和混料设计之间的联系. 文献中提供了许多建模方法, 如多项式回归模型、Kriging 模型、贝叶斯方法、神经网络以及诸多变量选择方法. 本书将简单介绍这些方法, 读者可参考 (Fang et al., 2006a) 了解其细节. 空间填充设计还有很多其他类型, 其中拉丁超立方体抽样应用最广泛. Santner 等 (2003) 和 Fang 等 (2006a) 给出了拉丁超立方体抽样的相关细节.

本书共 8 章. 第 1 章给出不同试验的简介及其相应的模型. 通过一个示例, 读者很容易理解均匀设计的核心思想及其具体方法. 本章也给出与试验设计有关的基本概念. 第 2 章考虑不同的均匀性度量并介绍其定义、计算表达式及性质, 并给出不同偏差的下界. 第 3 章和第 4 章介绍均匀设计的各种构造方法, 其中第 3 章侧重于确定性方法, 而第 4 章考虑随机优化算法. 第 5 章简单讨论不同的建模方

法. 均匀性不仅在构造均匀设计时扮演重要角色, 而且在其他类型的设计, 如因子设计、区组设计与超饱和设计中均可发挥重要作用. 第 6 章和第 7 章详细展示均匀性的重要意义. 第 8 章介绍混料试验的设计与建模.

本书可作为统计学与相关专业的研究生教材或工程技术人员的参考书, 我们已在多届本科生和研究生中教过本书的部分内容.

我们衷心感谢对均匀设计的发展做出重要贡献的合作者们, 他们包括中国科学院王元院士、伊利诺伊理工学院 Fred Hickernell 教授、宾州州立大学林共进教授、印度加尔各答管理学院 Rahul Mukerjee 教授、德国 Justus-Liebig-Universität Giessen 的 Peter Winker 教授、纽约州立大学水牛城分校马长兴教授、加州大学洛杉矶分校徐洪泉教授和印度 Visva-Bharati 大学 Kashinath Chatterjee 教授. 同时, 感谢下面各位教授的高效合作和鼓励, 其包括杨振海、张润楚、殷剑兴、潘建新、李润泽、陈令由、岳荣先、谢民育、唐煜、葛根年、梁逸曾、E. Liski、田国梁、宁建辉、杨建峰、孙法省、张爱军、欧祖军和 A. Elsawah 等. 我们特别感谢 K. Chatterjee 教授花了大量时间阅读本书英文版的初稿并给出诸多有益的建议.

本书的第一作者感谢香港 UGC 研究基金、北京师范大学–香港浸会大学联合国际学院研究基金 (资助号：R201409) 和珠海市优势学科基金的资助. 第二作者感谢国家自然科学基金 (11431006 和 11771220)、国家 “万人计划” 项目、天津市 “高层次创新创业团队” 基金和天津市 “131” 创新型人才团队基金的资助. 第三作者感谢国家自然科学基金 (11271147 和 11871237) 和华中师范大学中央高校基本业务费 (CCNU16 A02012 和 CCNU16JYKX013) 的资助. 第四作者感谢国家自然科学基金 (11471229 和 11871288) 和四川大学优秀青年基金 (2013SCU04A43) 的资助.

本书作者按姓名拼音排序.

方开泰
北京师范大学–香港浸会大学联合国际学院
中国科学院应用数学所

刘民千
南开大学统计与数据科学学院

覃　红
华中师范大学数学与统计学学院
中南财经政法大学统计与数学学院

周永道
南开大学统计与数据科学学院

2019 年 9 月

目 录

第 1 章　引　　言

试验设计是统计学的一个重要分支. 本章主要介绍各个领域的试验并介绍它们的重要性、目的、试验的类型、统计模型和相关的设计. 1.1 节给出了不同目的和特征的几种试验. 该节还讨论了两种流行的试验: ① 实体试验; ② 计算机试验. 试验设计中的一些基本术语在 1.2 节中介绍. 1.3 节介绍基于不同统计模型的各种试验设计的类型, 其中包括基于 ANOVA 模型的部分因子设计、基于线性回归模型的最优设计和模型未知 (或非参数回归模型) 下的均匀设计. 评价部分因子设计有很多准则, 其中最小低阶混杂准则被广泛应用, 1.4 节介绍该准则及其推广形式. 1.5 节介绍均匀设计在多因子试验中的实施过程. 建议读者仔细阅读本章, 以便能够很好地理解均匀设计的方法和本书随后的内容.

1.1　科 学 试 验

科学试验在探索大自然中非常重要, 几乎随处可见, 通常其目的是探知一个特定过程或系统. 科学试验经常被运用在农业、工业、自然科学和高新技术领域. 在工业和化工领域中, 试验的目的常常有:

1) 增加产量;

2) 提高产品的质量, 如降低变异性和增加可靠性;

3) 减少研发时间;

4) 降低总成本.

在自然科学和高新技术领域中, 科学试验的目的可能包括:

1) 评估材料的备选方案;

2) 筛选设计参数;

3) 确定对产品性能有关键影响的参数取值;

4) 评估因子效应和因子交互效应;

5) 探索因子与响应之间的联系.

当试验存在随机误差时, 对于一个具体的试验如何找到好的设计, 是统计学中的一个重要的研究领域. 试验的设计和建模是统计学的一个分支, 已经在科学和新技术的发展中发挥着重要的作用.

1.1.1　科学试验的例子

下面给出几个关于科学试验的例子, 这里忽略这些试验的具体细节, 使得读者专注于问题和后面要介绍的相关方法.

例 1.1　在一个化学试验中, 试验者希望通过试验探索某化学材料在合成物中的百分数(x)与其合成物强度(y)之间的关系. 假设如图 1.1 所示, 其内在的关系是未知的, 如何设计试验来找到一个近似模型 (或初始模型)来描述它们的关系? 在这个试验里, 化学材料的百分数被称为一个因子, 强度称为一个响应. 对于这个试验, 一个自然的想法是选择合成物的几个百分数, 例如, $x_1, \cdots, x_q$, 然后取这些值进行试验并测量相应的强度, 记为 $y_1, \cdots, y_q$. 各种建模的技术用来分析数据 $(x_1, y_1), \cdots, (x_q, y_q)$, 从而获得不同的模型. 试验者在这些模型中如何选择一个合适的试验点 $x_1, \cdots, x_q$ 和建模技术是非常重要的问题.

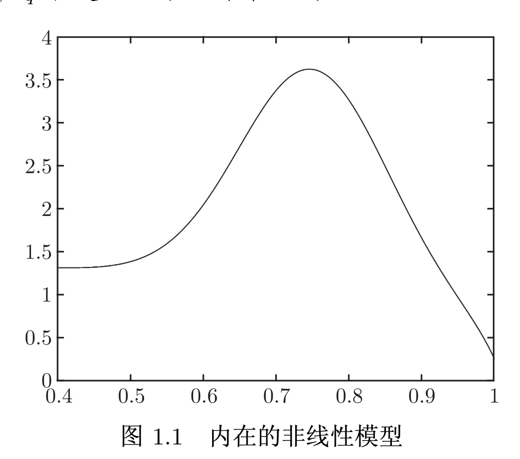


图 1.1　内在的非线性模型

例 1.2　这个试验是本书的第一作者于 1972 年在南京进行化工试验时遇到的一个典型问题. 为了提高某化工产品的产量 (y), 该试验研究了三个可控变量, 它们分别是

$$A: \text{反应温度;}$$
$$B: \text{反应容许时间;}$$
$$C: \text{碱浓度.}$$

这三个变量的变化区间分别为

$$A: 80-90℃;\quad B: 90-150\,\text{min};\quad C: 5\%-7\%.$$

根据经验可知: 这三个因子间没有交互效应, 并且产量 y 与这些因子间可能存在非线性的联系. 如何为这个试验进行设计? 在文献中有人推荐运用多因子试验. 1.3.1 节和 1.3.2 节将会介绍这类设计.

例 1.3 这个例子是 Fang 和 Chan(2006) 介绍的一个实际例子. 加速应力测试是研究系统寿命的重要方法. 随着技术的进步, 产品的寿命也会提高, 同时当新产品迅速涌现时, 它们的寿命周期将减少. 在下一代更新产品出现前, 制造商需要快速地确定新产品的寿命并投入市场. 在正常的操作条件下, 通过测试来确定产品的寿命在很多情形下是不可行的. 因此, 通常利用加速应力测试方法在高强度物理条件下来测试产品的寿命, 并利用寿命模型来推测产品失效的中位数时间. 基于物理或统计方面的考虑, 许多不同的模型, 例如, Arrhenius 模型、逆幂律模型、比例风险模型等, 被提出并用来模拟测试数据. 读者可查阅 (Elsayed, 1996) 来了解加速应力测试方法. 现需要一个试验设计来选择加速应力测试的环境参数. 考虑三个因子: 电压 V (伏特)、温度 T (开氏, 开尔文温标的计量单位) 和相对湿度 H (%). 响应 t 是失效的中位数时间, 其定义为

$$t = aV^{-b}e^{c/T}e^{-dH},$$

其中 a, b, c, d 是已知常数. 通过加速应力测试, 一个电子装置在正常的操作条件下的失效中位数时间 t_0 就可以被确定. 试验者希望通过试验来探索其模型, 并且在该研究中采用均匀设计进行试验.

例 1.4 在一个关于环境问题的研究中, 试验者希望对有毒的化学物质和它们的交互效应进行一个定量的风险评估. 六种化学物质分别为: 镉、铜、锌、镍、铬和铅. 为了判断化学物质浓度对毒性的影响, 试验者在试验中改变每一种化学物质的浓度. 但是响应和六种化学物质浓度之间内在的模型是未知的. 试验者想通过一个试验找到一个近似模型. 很明显, 每一种化学物质浓度的范围是非常重要的. 对于每一种化学物质, 试验者可以选择下面的浓度:

$$0.01, 0.05, 0.1, 0.2, 0.4, 0.8, 1, 2, 4, 5, 8, 10, 12, 14, 16, 18, 20.$$

考虑到每种化学物质都有 17 个水平, 六种化学物质共有 $17^6 = 24137569$ 种浓度组合, 超过了 2400 万组合! 因此, 对每一种浓度组合都进行试验是不可能的. 一个好的试验设计可以选取少数具有代表性的浓度组合进行试验并可以获得一些有用的结果. 有关这类试验设计和建模问题的详细介绍可参见 (Fang and Wang, 1994; Fang et al., 2006a).

例 1.5 在化学和化学工程中, 可逆化学反应是一类重要的基础反应. 传统上, 根据化学反应的特性, 化学家使用确定性的方法可以获得动态的速率常数. 化学家常使用的方法是在反应中使用大量过剩的反应物, 以便让它们浓度的变化可以忽略, 因而可以获得反应物和产品之间的简单关系. 如下 11 个微分方程所构成的线性系统可以用来模拟化学动力学:

$$h_j(x,t) = g_j(\eta, x, t), \quad j = 1, \cdots, 11, \tag{1.1}$$

其中 x 是速率常数的集合, 即系统的输入. 对于任意的输入 x, 利用这 11 个微分方程, 在每一次反应时间 7×10^{-4}s 中, 就可以产生 5 种化学物种的浓度, 从而得到 (1.1) 的一个解. 试验者感兴趣的是希望找到一个比真实模型简单得多的近似模型. 对于可逆化学反应的动力学试验, Atkinson 等 (1998) 讨论了运用 D 最优设计的可行性.

例 1.6　许多产品是由两个或多个材料混合形成的. 比如, 制作一个咖啡面包, 材料有: X_1 (面粉)、X_2 (水)、X_3 (糖)、X_4 (植物起酥油)、X_5 (椰子薄片)、X_6 (盐)、X_7 (酵母)、X_8 (乳化剂)、X_9 (丙酸钙)、X_{10} (咖啡粉) 和 X_{11} (液体香精). 对于每一种成分, 如何选择一个合适的百分比? 这需要烹饪师根据其丰富的经验来选择. 假如烹饪师没有任何经验, 混料试验设计是非常有用的. 注意到上述 11 种成分必须满足 $x_1+\cdots+x_{11}=1$ 和 $x_i\geqslant 0, i=1,\cdots,11$. 这些限制意味着这类试验设计是不同于前面介绍的其他试验, 这种试验称为混料试验. 此外, 这里提到的限制也可能是某一种成分量的上界和下界, 即 $a_i\leqslant x_i\leqslant b_i, i=1,\cdots,s$. 例如, 在咖啡面包制作试验中, 水和面粉是主要成分, 盐量低于 1%, 其他成分的用量也是非常小, 这样的试验称为带限制的混料试验, 这类试验在新型材料的研发中非常有用.

一个具有 n 次试验、s 个成分的混料设计可以看作试验域 T^s 中的 n 个点的集合, 这里

$$T^s=\{(x_1,\cdots,x_s):x_j\geqslant 0,\ j=1,\cdots,s,\ x_1+\cdots+x_s=1\}. \tag{1.2}$$

由于 $x_1+\cdots+x_s=1$, 因此找到一个混料试验设计与找到一个对因子没有任何限制的试验设计完全不同.

例 1.7 (机器人手臂)　机器人是由计算机程序控制的电机机械. 机器人技术的一个新的分支包括了设计、结构、操作和计算机系统. 计算机手臂的运动轨迹常用于解释计算机试验的一个例子 (参考文献 (Fang et al., 2006a)). 考虑一个具有 m 段的机器人手臂. 手臂的肩部固定在平面 (u,v) 中的原点. 手臂的第 j 段有长度 $L_j, j=1,\cdots,m$. 相对于平面的水平坐标轴, 第 1 段手臂在角 θ_1 的方向. 当 $k=2,\cdots,m$ 时, 第 k 段手臂与第 $k-1$ 段手臂的夹角是 θ_k. 最后一段手臂的末端在

$$\begin{cases} u=\displaystyle\sum_{j=1}^{m}L_j\cos\left(\sum_{k=1}^{j}\theta_k\right), \\ v=\displaystyle\sum_{j=1}^{m}L_j\sin\left(\sum_{k=1}^{j}\theta_k\right). \end{cases} \tag{1.3}$$

响应 y 是计算机手臂的末端到起始点的距离, $y=\sqrt{u^2+v^2}$ 可表示成 $2m$ 个变量

$\theta_j \in [0, 2\pi]$ 和 $L_j \in [0, 1]$ 的函数. 对于三段的计算机手臂, Ho (2001) 给出了一个近似模型 $y = g(\theta_1, \theta_2, \theta_3, L_1, L_2, L_3)$.

1.1.2 试验的特征

在前面的小节中, 我们列举了几个试验, 每一个试验有其自己的目的和特征. 例 1.1 想探究化学材料在合成物中的百分数 (x) 与其合成物强度 (y) 之间的非线性关系, 这是一个单因子试验且具有模型不确定性. 例 1.2 是一个多因子试验并需要一个有效的试验设计来探讨响应和因子之间的关系, 且其中给出的一些先验知识 (因子之间无交互效应) 是非常重要的. 例 1.3 研究响应 t 和三个因子 V, T, H 之间的非线性关系, 并选择加速应力测试的环境参数. 例 1.4 的目标是至关重要的, 但是响应和六种化学物质浓度之间的关系非常复杂并且内在的模型是未知的. 例 1.5 来自化学和化学工程. Atkinson 等 (1998) 讨论了当内在的模型已知时, 将 D 最优设计运用于可逆化学反应的动力学的可行性. 但是, 当模型变化时, D 最优设计是不稳健的. 对于这种化学反应, Xu 等 (2000) 考虑了三种不同的设计并比较它们的表现. 他们断定对于这种试验, 均匀设计是稳健的. 许多产品是用几种成分混合形成的, 混料试验在许多领域起着非常重要的作用, 像化学工程、橡胶工业、材料和制药工程. 例 1.6 就是其中一个例子. 例 1.7 和计算机试验相关, 内在的模型是已知的, 但很复杂. 试验者希望找到一个近似模型替代真实模型, 计算机试验的设计和建模是一个发展迅速的领域, 并被广泛运用于系统工程.

一个好的试验设计应该是用尽可能少的试验次数来获取尽可能多的信息. 试验设计是统计学的一个分支并具有悠久的历史. 试验设计在科学和新技术发展中起着非常重要的作用, 特别是在高科技的发展中, 它包含了丰富的方法和各种设计. 对于各种类型设计的综述可参考由 S. Ghosh 和 C. R. Rao 所编辑的*Handbook of Statistics*, Vol. 13, 由 R. Khattree 和 C. R. Rao 编辑的*Handbook of Statistics*, Vol. 22: *Statistics in Industry* 与由 A. Dean, M. Morris, J. Stufken 和 D. Bingham 编辑的 *Handbook of Design and Analysis of Experiments*. 在自然科学发展史上, 一些自然规律就是通过精心设计的试验被发现或证实的, 通过这些试验所获得的信息和推断可能导致一个新的科学分支的诞生. 生命科学的发展常以这种方式出现, 众所周知的一个经典例子是孟德尔基因遗传模型.

试验设计在多因子试验中是非常有用的. 从上面的例子我们发现, 试验特别是在高科技中的试验, 具有如下一个或多个特征:

(1) 多因子;

(2) 响应和因子之间的非线性关系;

(3) 试验区域非常大以致因子的水平数不能太少;

(4) 内在的模型是未知的;

(5) 试验可以通过计算机进行模拟.

我们需要一些新的设计来处理具有上述复杂性的试验. 已有许多试验设计能满足上述要求, *均匀试验设计*或*均匀设计*就是其中之一.

1.1.3 试验类型

基于各种考虑, 试验可以从不同的角度来分类. 例如,

I. 因子

(1) 单因子试验;

(2) 多因子试验.

II. 响应个数

(1) 单个响应;

(2) 多重响应;

(3) 泛函响应 (如声音、指纹、心电图).

III. 约束条件

(1) 无约束试验: 每个因子的水平选择是独立的;

(2) 约束试验: 每个因子的水平选择是依赖于其他因子的选择, 如混料试验.

IV. 响应和因子之间的内在的模型

(1) 已知;

(2) 未知.

V. 操作环境

(1) 实体试验;

(2) 计算机试验.

我们对最后一类试验分类进行如下解释.

A. 实体试验

从传统意义上来讲, 试验是在实验室、工厂或田间进行操作的. 这样的试验叫做*实体试验*或*实际试验*, 试验者自己实施试验. 在实体试验中一定存在着随机误差, 这使得在相同的试验环境下得到不同的输出. 随机误差的存在使得数据分析和建模的过程变得复杂. 因而, 试验者在试验中可能选择一个或几个因子以便于研究输入与输出之间的关系, 或提出一些强有力的统计试验设计方法.

设计试验的统计方法通常是基于响应和因子之间的统计模型. 一个好的设计对于探察内在的统计模型是最优的. 对于实体试验存在着许多设计, 包括基于 ANOVA 模型的*部分因子设计*和基于回归模型的*最优回归设计*(简称最优设计), 这些设计在实践中都非常受欢迎. 其相应的模型包含了一些未知参数, 如主效应、交互效应、回归参数和随机误差的方差. 一个好的设计在某种意义下用较小甚至最小

的方差–协方差矩阵对这些参数给出无偏估计. 正交设计和各种最优设计就是这样的常用设计. 实体试验中常用的概念、正交表和最优设计的基本知识将在后面的小节里具体介绍.

正交设计和最优回归设计都假定内在的模型是已知, 模型中的参数是未知的. 但是在大部分试验中, 内在模型是未知的. 基于这些试验, 试验者希望找到一个近似的模型. 均匀设计是最受欢迎的设计之一.

B. 计算机试验

计算机试验或计算机模拟被广泛地运用于工业、系统工程和其他领域中的物理现象的研究, 因为许多物理过程/现象用常用的试验方法去研究是困难的, 甚至是不可能的. 计算机模型被用于描述在科学和工程中遇到的复杂的物理现象. 这些现象常常受控于有关的方程组, 包括线性、非线性、常微分和偏微分方程组. 解这些方程组通常太困难, 在短时间内不能求出解, 但通过计算机建模程序可以求出. 方程的数量和复杂性, 这些程序可能需要运行很长时间, 这使得用它们做广泛的科学研究是困难的. Santner 等 (2003) 指出 "通过常规的方法研究许多物理过程是困难的甚至是不可能的. 随着计算能力的增加, 通过复杂的计算机编码对这些过程中一些过程进行建模变为可能".

在过去的几十年中, *计算机试验*或*模拟试验*在统计和工程中成为一个热门的话题, 计算机试验中内在的模型是确定的, 但是它太复杂而不能被处理和分析. 计算机试验的目标之一是首先找到一个近似模型, 该模型比真实模型简单. 模拟试验是通过在计算机上模拟过程来研究内在的关系. 模拟试验的内在模型是确定的并被给定, 但是需要考虑输入的误差.

计算机试验作为人工智能可以在模拟物理环境方面发挥其作用. 假如因为某些原因, 一些试验不能进行实体试验, 但计算机试验可以使得这些试验几乎可以实现. 许多科学研究涉及使用如下数学模型对复杂的物理现象进行建模:

$$y = f(x_1, \cdots, x_s) = f(\boldsymbol{x}), \quad \boldsymbol{x} = (x_1, \cdots, x_s)^{\mathrm{T}} \in T, \tag{1.4}$$

其中 $\boldsymbol{x}$ 是输入变量, $\boldsymbol{x}^{\mathrm{T}}$ 是 $\boldsymbol{x}$ 的转置, y 是输出变量, 函数 f 可能不是一个解析式, T 是输入变量空间. 模型 (1.4) 可看作方程组的一个解, 包括线性、非线性、常微分或偏微分方程组. 对这些方程组, 要得到一个解析解通常是不可能的或用常规试验方法去研究也是不可能的. 工程师或科学家利用模型去完成各种任务和在不同输入变量设置下, 通过询问模型预测系统行为而进行决策. 因此, 在科学研究和工程设计中, 模型起着关键性的作用. 但是, 在许多情况下, 为了充分了解响应变量 y 和输入变量 x_j 之间的关系进行实体试验是非常昂贵和耗时的. 作为一种替换, 科学家和工程师使用计算机模拟研究这种关系. 因而, 近似模型在研究复杂的物理现

象变得非常重要. 计算机试验目标之一是首先在计算机上通过模拟研究找到一个好的近似模型, 该模型比真实模型简单 (图 1.2). 由于方程组的数量和复杂性, 这些程序需要进行特殊的设计.

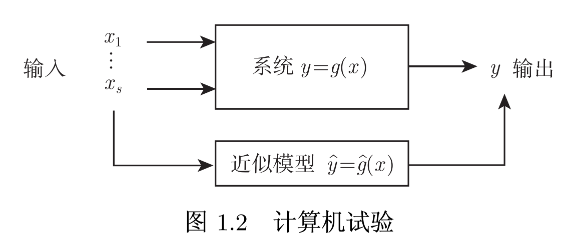


图 1.2　计算机试验

1.2　基 本 术 语

本节介绍在试验设计中用到的一些基本术语. 先看一个典型的例子.

例 1.8　在一个化学试验中, 为了增加产量, 三个可控变量被研究. 它们是

A (或 x_1): 原材料的类型;

B (或 x_2): 吡啶的数量;

C (或 x_3): 反应持续的时间.

这些变量的范围分别是 $\{m_1, m_2, m_3\}$, $[10, 28]$ 和 $[0.5, 3.5]$. 在试验中, 每个变量取值的范围如下:

A: m_1, m_2, m_3;

B (毫升): 10, 19, 28;

C (小时): 0.5, 1.5, 2.5, 3.5.

因子　指的是试验中感兴趣的可控变量. 为了方便, 因子记为 $A, B, \cdots$ 或 x_1, $x_2, \cdots$. 例 1.8 中, 原材料的类型、吡啶的数量和反应持续的时间是因子. 因子可能是定量的或定性的. 定量因子可以通过数值度量或位于一个区间中, 例如, 温度、压力、两种原材料的比率、反应时间等等. 定性因子是通过分类进行度量, 像教育程度、材料类型、性别等等. 定性因子也称为分类因子或指示因子. 在例 1.8 中, 因子 A 是定性的, 因子 B 和 C 是定量的. 在计算机试验中, 这些因子被称为输入变量.

环境变量和噪扰变量　环境变量指的是在试验中不需要研究的变量, 这些变量被限定在事先确定的值. 环境变量不被看作因子. 根据例 1.8, 其他化学物质的数量、压力等等都是环境变量.

不可控的变量都看作噪扰变量或噪声变量. 在大部分农田试验中, 天气是不可控的, 可能产生一些噪扰变量. 在许多实体试验中, 季节被看作噪扰变量.

试验域 指的是因子取值的空间. 在例 1.8 中, 试验域是*笛卡儿积*$\{m_1, m_2, m_3\} \times [10, 28] \times [0.5, 3.5]$. 用 $\mathcal{X}$ 表示试验域, 它是 R^s 的一个子集, 其中 s 是因子数量. 在新产品或新程序的发展中, 选择一个较大的试验域是比较好的, 这样会有更多的选择, 找到理想的结果. 在工厂生产环境下的试验, 为了安全起见, 推荐小的试验域.

水平 指的是因子被测试时所取的具体数值. 因子 A 的水平数记为 $A_1, A_2, \cdots$. 在例 1.8 中, 反应时间的水平为 $C_1 = 0.5, C_2 = 1.5, C_3 = 2.5, C_4 = 3.5$, 原材料类型的水平为 $A_1 = m_1, A_2 = m_2, A_3 = m_3$. 通常, 试验者首先选择一个试验域, 然后对每个因子进行水平选择.

水平组合 指的是因子水平的一个可能组合. 根据例 1.8, $(m_2, 10, 2.5)$ 和 $(m_1, 28, 0.5)$ 是两个水平组合. 一个水平组合也称为一个*处理组合*, 可以看作试验域 $\mathcal{X}$ 中的一个点.

试验 指的是在试验环境中, 一个水平组合的实施. 在本书中, 用 “n” 表示试验的次数.

试验设计 指的是水平组合的集合或在试验域 $\mathcal{X}$ 中的一个点集. 具有 n 次试验和 s 个因子的设计记作 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$, 其中 $\boldsymbol{x}_j \in \mathcal{X}$ 或 $n \times s$ 矩阵 $\boldsymbol{U} = (u_{ij})$, $\boldsymbol{U}$ 中的行分别是 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 的转置. 为了简单起见, 有时设计 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 可以看作矩阵 $\boldsymbol{U}$.

响应 指的是取决于试验目的的一次试验的结果. 响应 (或试验的输出) 可能是定量的或定性的, 常用 y 表示. 例 1.8 中, 试验的响应是产量. 有时一次试验对应于几个响应. 这样的试验称为*多重响应试验*. 例如, 在化学、生物和医药研究中, 响应可能是一条曲线, 像化学指纹或 DNA 指纹. 曲线响应称为*泛函响应*.

随机误差 指的是由于试验中的不可控变量的影响而产生的响应变异. 一般来说, 因为随机误差, 对两次相同的水平组合进行试验, 试验的结果不尽相同. 实体试验存在随机误差, 在相同的试验环境下, 重复去做试验叫做重复试验. 由于随机误差, 重复试验得到的结果可能不同. 在大多数试验中, 随机误差通常被假定服从*正态分布*$N(0, \sigma^2)$. 方差 σ^2 用来度量随机误差的大小. 随机误差的存在使得设计和建模更加复杂.

因子的交互效应 假设 A 和 B 是试验中考虑的两个因子. 假如因子 A 对响应的影响不依赖于因子 B 所取的水平, 因子 B 对响应的影响也不依赖于因子 A 所取的水平, 则认为因子 A 和 B 之间不存在交互效应. 否则我们就认为因子 A 和 B 之间存在交互效应. 类似地, 可以定义更多因子之间的交互效应. 怎样定义和估计因子间的交互效应是一个重要问题. 关于这方面的讨论见 1.3 节.

下面将给出在本书中用到的一些有用的代数概念.

Hadamard 矩阵 n 阶 Hadamard 矩阵 $\boldsymbol{H}$ 是一个 $n \times n$ 矩阵, 其中的元素为 1 或 -1, 并满足 $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}} = n\boldsymbol{I}$. 很明显, n 必须是 1, 2 或 4 的倍数. Hadamard 矩阵

在编码理论和试验设计中起着非常重要的作用.

Hadamard 积或点积 两个 $n \times m$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 和 $\boldsymbol{B} = (b_{ij})$ 的 Hadamard 积是一个 $n \times m$ 矩阵, 该矩阵中每一个有序对 (i, j) 的元素为 $a_{ij}b_{ij}$.

$$\boldsymbol{A} \cdot \boldsymbol{B} = \begin{pmatrix} a_{11}b_{11} & \cdots & a_{1m}b_{1m} \\ \vdots & & \vdots \\ a_{n1}b_{n1} & \cdots & a_{nm}b_{nm} \end{pmatrix}. \tag{1.5}$$

Hadamard 积在文献中也称为*点积*.

Hamming 距离 具有相同长度的两个字符串之间的 Hamming 距离是相应位置取不同字符的位置数. 例如, 两个字符串 (a, f, t, e, r) 和 (f, f, b, r, r) 有 Hamming 距离 3; 在 $L_9(3^4)$ 中的两次试验 $(2, 1, 2, 3)$ 和 $(2, 2, 3, 1)$ 有 Hamming 距离 3.

Kronecker 乘积 $n \times p$ 的矩阵 $\boldsymbol{A} = (a_{ij})$ 和 $m \times q$ 的矩阵 $\boldsymbol{B} = (b_{kl})$ 的 Kronecker 乘积是一个 $nm \times pq$ 矩阵, 该矩阵被定义为

$$\boldsymbol{A} \otimes \boldsymbol{B} = (a_{ij}\boldsymbol{B}) = \begin{pmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1m}\boldsymbol{B} \\ \vdots & & \vdots \\ a_{n1}\boldsymbol{B} & \cdots & a_{nm}\boldsymbol{B} \end{pmatrix}. \tag{1.6}$$

笛卡儿积 数学中, 笛卡儿积是一个数学运算, 该运算让多个集合变为一个集合. 例如, $\mathcal{A} = \{a_1, \cdots, a_m\}$ 和 $\mathcal{B} = \{b_1, \cdots, b_n\}$ 的笛卡儿积是集合

$$\mathcal{A} \times \mathcal{B} = \{(a, b) : a \in \mathcal{A}, b \in \mathcal{B}\}. \tag{1.7}$$

更一般地, n 个集合的笛卡儿积也称为 n 重笛卡儿积, 可以表示为一个 n 维数组, 其中每一个元素是一个 n 元组.

按模运算 在本书中, 按模运算具有特殊的意义. 设 m 和 n 是两个正整数. 很容易发现 m 可以表示为

$$m = qn + r, \tag{1.8}$$

其中非负整数 $q \leqslant m/n$, r 是 m 除以 n 的余数. 当 $m \leqslant n$, 有 $q = 0$ 和 $r = m$; 当 $m > n$, q 是一个正整数, $0 \leqslant r < n$. 定义 $m(\bmod\ n) = r$. 很明显, $m(\bmod\ n)$ 的范围是 $\{0, 1, \cdots, n-1\}$. 在试验设计中, 按模运算有时被修改为当 $r > 0$ 时, $m(\widetilde{\bmod}\ n) = r$; 当 $r = 0$ 时, $m(\widetilde{\bmod}\ n) = n$.

1.3 统 计 模 型

每一个试验都有自己的目标. 一个带有某些未知参数的特定统计模型可以用来

拟合试验中因子和响应之间的关系. 正如前面所提到的两种试验类型: ① 内在模型是已知的, 但包括一些未知的参数; ② 真实模型是未知的.

当内在模型是已知的, 一些最优准则被提出并用来评价参数的估计. 我们希望找到一个设计, 这个设计所获得的数据在最优准则下能对参数进行最好的估计. 基于具体的目标模型, 许多文献已给出相应的各种统计设计. 对于这些设计的详细介绍可参考由 Ghosh 和 Rao (1996) 主编的*Handbook of Statistics*, Vol. 13 与由 Dean 等 (2015) 主编的*Handbook of Design and Analysis of Experiments*.

我们使用下面的例子解释关于不同类型的统计模型的运用.

例 1.1(续) 假定试验者根据其专业知识选定试验区域为 $\mathcal{X} = [0.4, 1]$ 和根据资金预算情况选取试验次数为 12 次. 为制订这个试验的设计有几种统计方法, 每一种方法都有自己的统计模型. 在下面的小节中将介绍基于因子设计的 ANOVA 分析模型、基于最优设计的回归模型以及受均匀设计启发的非参数回归模型.

1.3.1 部分因子设计和 ANOVA 模型

部分因子设计是水平组合的集合, 试验的主要目的之一是估计主效应和因子交互效应. 因子设计被广泛运用于各种类型的试验, 这里给出一个简单的介绍.

A. 单因子试验

在例 1.1 中, 试验者使用因子设计, 选择化学材料在合成物中的百分数 $x_1, \cdots, x_q$, 即水平. 然后对于每一个 x_i, 重复 n_i 次试验, 相应的响应记为 $y_{i1}, \cdots, y_{in_i}$. 图 1.3 给出了当 $q = 4$, $n_1 = n_2 = n_3 = n_4 = 3$ 时的试验结果. 考虑如下因子设计统计模型:

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \cdots, q, \ j = 1, \cdots, n_i, \tag{1.9}$$

其中 μ_i 是 y 在 $x = x_i$ 时的真实值, 记为 $y(x_i)$, ε_{ij} 是随机误差, 它们独立同分布且服从正态分布 $N(0, \sigma^2)$, $\sigma^2 > 0$ 未知. 假设 μ 是 y 在 $x_1, \cdots, x_q$ 上的总均值. 均值 μ_i 可分解为 $\mu_i = \alpha_i + \mu$, 其中 α_i 叫做 y 在 x_i 上的主效应并且满足 $\alpha_1 + \cdots + \alpha_q = 0$. 于是模型 (1.9) 可以表示为

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1, \cdots, q, \ j = 1, \cdots, n_i. \tag{1.10}$$

主效应衡量了因子 A 对响应的影响. 假如某些 $\alpha_i \neq 0$, 这说明因子 A 对响应 y 有显著影响. 用最小二乘法可以获得 $\mu, \alpha_1, \cdots, \alpha_{q-1}$ 的估计, 方差分析可以给出 σ^2 的无偏估计以及 F-检验. 因而, 模型 (1.10) 称为 ANOVA 模型.

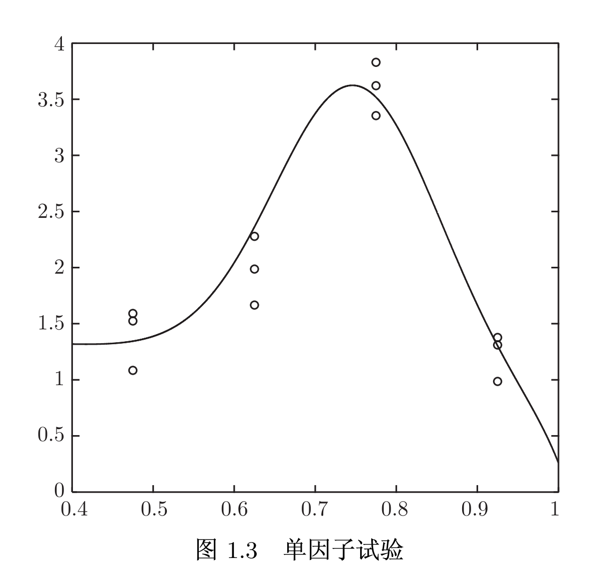


图 1.3 单因子试验

B. **二因子试验**

对于二因子试验, 它的方差分析模型包括总均值、主效应和因子之间的交互效应. 假定在二因子试验中, 因子 A 选 q_1 个水平 $A_1,\cdots,A_{q_1}$, 因子 B 选 q_2 个水平 $B_1,\cdots,B_{q_2}$. 对于每一个水平组合, 重复 r $(r\geqslant 1)$ 次. 该试验的统计模型可表示为

$$\begin{aligned} &y_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk},\\ &i=1,\cdots,q_1,\ j=1,\cdots,q_2,\ k=1,\cdots,r, \end{aligned} \tag{1.11}$$

其中

y_{ijk}: 在水平组合 A_iB_j 上的第 k 次响应值;

μ: 总均值;

α_i: 因子 A 在水平 A_i 上的主效应;

β_j: 因子 B 在水平 B_j 上的主效应;

$(\alpha\beta)_{ij}$: 因子 A 和 B 在水平组合 A_iB_j 上的交互效应;

ε_{ijk}: 在 A_iB_j 上的第 k 次试验中随机误差. 通常, 假定 $\{\varepsilon_{ijk}\}$ 独立同分布且服从正态分布 $\varepsilon_{ijk}\sim N(0,\sigma^2)$, 这里 σ^2 是未知的.

模型中的主效应和交互效应满足

$$\alpha_1+\cdots+\alpha_{q_1}=0;\quad \beta_1+\cdots+\beta_{q_2}=0;$$

$$\sum_{i=1}^{q_1}(\alpha\beta)_{ij}=0,\ j=1,\cdots,q_2;\quad \sum_{j=1}^{q_2}(\alpha\beta)_{ij}=0,\ i=1,\cdots,q_1.$$

因此, A 的线性无关的主效应个数是 q_1-1 个, B 的线性无关的主效应个数是 q_2-1 个, $A\times B$ 的线性无关的交互效应个数是 $(q_1-1)(q_2-1)$ 个. 这里的 q_1-1, q_2-1 和 $(q_1-1)(q_2-1)$ 在 ANOVA 中分别叫做 A, B 和 $A\times B$ 的*自由度*. 模型 (1.11) 可以表示为

$$y_{ijk}=\mu_{ij}+\varepsilon_{ijk},$$

其中 μ_{ij} 是 y 在水平组合 A_iB_j 上的均值. 将 μ_{ij} 改写为

$$\mu_{ij}=\mu+(\mu_{i\cdot}-\mu)+(\mu_{\cdot j}-\mu)+(\mu_{ij}-\mu_{i\cdot}-\mu_{\cdot j}+\mu), \tag{1.12}$$

其中 $\mu_{i\cdot}$ 是 y 在水平 A_i 上的均值, $\mu_{\cdot j}$ 是 y 在水平 B_j 上的均值. 上面的划分表明

$$\alpha_i=\mu_{i\cdot}-\mu,\quad \beta_j=\mu_{\cdot j}-\mu,\quad (\alpha\beta)_{ij}=\mu_{ij}-\mu_{i\cdot}-\mu_{\cdot j}+\mu. \tag{1.13}$$

这些参数的最小二乘估计分别为

$$\begin{aligned}\hat{\mu}&=\bar{y},\\ \hat{\alpha}_i&=\bar{y}_{i\cdot}-\bar{y},\quad \hat{\mu}_{i\cdot}=\bar{y}_{i\cdot},\\ \hat{\beta}_j&=\bar{y}_{\cdot j}-\bar{y},\quad \hat{\mu}_{\cdot j}=\bar{y}_{\cdot j},\\ (\widehat{\alpha\beta})_{ij}&=\bar{y}_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y},\quad \hat{\mu}_{ij}=\bar{y}_{ij},\end{aligned}$$

其中

$\bar{y}_{ij}$: 在水平组合 A_iB_j 上的平均响应;
$\bar{y}_{i\cdot}$: 在水平 A_i 上的平均响应;
$\bar{y}_{\cdot j}$: 在水平 B_j 上的平均响应;
$\bar{y}$: 总的平均响应.
同时, 可以获得 ε_{ijk} 的估计, 即*残差* $\hat{\varepsilon}_{ijk}=y_{ijk}-\bar{y}_{ij}$.

C. 多因子试验

具有三个因子 A, B 和 C 的试验包括 A,B 和 C 的主效应, $A\times B$, $A\times C$ 和 $B\times C$ 的两因子交互效应, $A\times B\times C$ 的三因子交互效应. 假定因子 A 有 q_1 个水平, 因子 B 有 q_2 个水平, 因子 C 有 q_3 个水平. 假设模型为

$$\begin{aligned}y_{ijkt}=&\mu+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}+(\alpha\beta\gamma)_{ijk}+\varepsilon_{ijkt},\\ &i=1,\cdots,q_1,\quad j=1,\cdots,q_2,\quad k=1,\cdots,q_3,\quad t=1,\cdots,r,\end{aligned} \tag{1.14}$$

其中 A, B 和 C 的主效应分别有 q_1-1, q_2-1 和 q_3-1 个自由度; 它们的二因子交互效应 $A\times B$, $A\times C$ 和 $B\times C$ 分别有 $(q_1-1)(q_2-1)$, $(q_1-1)(q_3-1)$ 和

$(q_2-1)(q_3-1)$ 个自由度; $A\times B\times C$ 三阶交互效应有 $(q_1-1)(q_2-1)(q_3-1)$ 个自由度. 假如我们需要估计上面所有的主效应和交互效应以及误差的方差 σ^2, 则试验次数至少为 $q_1q_2q_3+1$. 一般地, 对于一个有 s 个因子、每个因子分别有 $q_1,\cdots,q_s$ 水平的试验, 假如需要考虑所有的主效应、二因子交互效应、三因子交互效应 $\cdots\cdots s$ 阶因子交互效应, 那么被估计参数的总个数是 $\prod_{i=1}^s q_i+1$. 当 $(s,q_1,\cdots,q_s)$ 增加时, 被估计的参数呈指数增加. 如果需要估计上述的所有未知参数, 那么我们推荐使用完全因子设计.

定义 1.1 一个完全因子设计或完全设计是所有水平组合出现相同次数的因子设计.

完全因子设计的优点是可以估计出所有因子的主效应和因子间的交互效应. 但是, 完全因子设计中参数的数量（主效应和交互效应）随着因子个数的增加呈指数递增. 例如, 一个试验有 6 个因子, 每个因子有 5 个水平, 则有 $5^6+1=15626$ 个未知参数需要估计, 包括估计 σ^2. 假如运用完全因子设计, 试验的最小次数为 15626 次. 另一方面, 有重复试验的试验次数应该是 15625 的倍数, 这样对于一个实体试验而言, 试验次数太大了! 表 1.1 给出了在各种 ANOVA 模型中未知参数的个数.

表 1.1 未知参数的个数

因子数 (s)	1	2	3	4	5	6	7	8
线性 $(q=2)$	2	4	8	16	32	64	128	256
二次 $(q=3)$	3	9	27	81	243	729	2187	6561
三次 $(q=4)$	4	16	64	256	1024	4096	16384	65536

1.3.2 部分因子设计

在多个因子的试验中, 需要包含所有未知参数的模型吗? 根据经验, 试验者认为因子试验的 ANOVA 模型应该满足下面两个原则:

稀疏原则 在因子设计中, 相对重要的主效应和交互效应的数量较小.

等级排序原则 低价效应比高阶效应更重要; 主效应比交互效应更重要; 同阶效应同等重要.

如果考虑到上述两个原则, 那么未知参数的个数会显著减少. 例如, 对于一个有 s 个因子、每个因子有 q 个水平的试验, 存在 $s(q-1)$ 个主效应和 $s(s-1)(q-1)^2/2$ 个二因子交互效应. 假如高阶交互效应可以忽略, 那么未知参数的个数从 q^s 减少到 $s(q-1)[1+(s-1)(q-1)/2]+1$. 假设某些二因子交互效应不显著, 则未知参数的个数会进一步减少. 基于这种考虑, 完全因子设计的一个子集被采用, 称这个子集

为部分因子设计(FFD). 关于部分因子设计的细节, 读者可参考 (Dey and Mukerjee, 1999). 在文献中常用的一些标记, 如 2^{s-k} 表示的是 s 个二水平因子、试验次数为 $n=2^{s-k}$ 的部分因子设计. 这个设计是完全因子设计 2^s 的 $1/2^k$ 部分设计. 类似地, 符号 3^{s-k} 表示的是 s 个三水平因子、试验次数为 $n=3^{s-k}$ 的部分因子设计.

对于 FFD 的使用, 试验者需要有一些关于哪些主效应和交互效应是显著的先验知识. 能估计模型所有未知参数的设计称为*可估的*. 对于后者, 一些主效应和交互效应可能存在混杂, 这种现象称为*混淆现象*.

好的部分因子设计是用有限的试验次数可以估计出所有的主效应和感兴趣的交互效应, 并且允许不重要的参数混杂. 最常见的部分因子设计是*正交表*.

定义 1.2 *一个 n 次试验、s 个有 q 个水平的因子、强度为 r 的正交表记为 $\mathrm{OA}(n,s,q,r)$, 它是一个 FFD, 确保它中间包含 n 次试验和 m $(m \leqslant r)$ 个因子且每个因子有 q 个水平的任意子设计是一个完全因子设计.*

强度为 2 的正交表被广泛运用于各个领域中的计划试验 (参见 (Hedayat et al., 1999)). 正交表可以表示为一个表. 一般地, 定义如下.

定义 1.3 *一个 $n \times s$ 的矩阵被称为正交设计表, 记为 $L_n(q_1 \times \cdots \times q_s)$, 它的第 j 列有 q_j 个不同的元素, 例如, $1,2,\cdots,q_j$, 并且它满足:*

(1) 每一列所有的元素出现相同的次数;

(2) 任意两列的所有元素组合出现的次数相同.

很显然, 条件 (2) 包含了条件 (1). 当某些 q_i 相同时, 正交设计表 $L_n(q_1\times\cdots\times q_s)$ 被记为 $L_n(q_1^{r_1} \times \cdots \times q_m^{r_m})$, 其中 $r_1+\cdots+r_m=s$. 假如所有的 $q_i=q$, 则正交设计表 $L_n(q_1 \times \cdots \times q_s)$ 记为 $L_n(q^s)$. 表 1.2 和表 1.3 给出了正交设计表的两个例子.

表 1.2 正交设计表 $L_8(2^7)$

编号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

表 1.3 正交设计表 $L_8(4 \times 2^4)$

编号	1	2	3	4	5
1	1	1	1	1	1
2	1	2	2	2	2
3	2	1	1	2	2
4	2	2	2	1	1
5	3	1	2	1	2
6	3	2	1	2	1
7	4	1	2	2	1
8	4	2	1	1	2

定义 1.4(饱和设计) 假如正交设计 $L_n(q^s)$ 中的 n, q 和 s 满足 $n-1=s(q-1)$, 这样的正交设计称为饱和的. 更一般地, 假如对于正交设计 $L_n(q_1 \times \cdots \times q_s)$, 它满足 $n-1=\sum_{j=1}^{s}(q_j-1)$, 那么称其为饱和的.

设计 $L_9(3^4)$, $L_8(2^7)$ 和 $L_8(4 \times 2^4)$ 是饱和的. 设计 $L_{18}(3^7)$ 是不饱和的, 因为 $(3-1) \times 7 = 14 < 17 = (18-1)$, 且设计 $L_{18}(2 \times 3^7)$ 也是不饱和的. 饱和设计不能被插入任意一列以致新的设计仍然是正交的.

均匀设计可看作模型未知的一种部分因子设计.

定义 1.5 (均匀设计表) 称一个 $n \times s$ 矩阵为均匀设计表, 记为 $U_n(q_1 \times \cdots \times q_s)$, 第 j 列的元素为 $1, 2, \cdots, q_j$, 假如它满足:

1) 每一列所有的元素出现相同的次数;

2) 由 $U_n(q_1 \times \cdots \times q_s)$ 确定的 n 个试验点在某种意义上均匀地分布在试验域中.

当某些 q_i 相同时, 均匀设计表 $U_n(q_1 \times \cdots \times q_s)$ 记为 $U_n(q_1^{r_1} \times \cdots \times q_m^{r_m})$, 其中 $r_1 + \cdots + r_m = s$. 假如所有的 $q_i = q$, 则均匀设计表 $U_n(q_1 \times \cdots \times q_s)$ 记为 $U_n(q^s)$.

均匀设计是建立在一个非参回归模型上的一种设计, 这里非参回归模型为

$$y = f(x_1, \cdots, x_s) + \varepsilon, \tag{1.15}$$

其中真实模型 $f(\cdot)$ 未知, $(x_1, \cdots, x_s)$ 是试验域中的试验点, ε 是随机误差. 均匀设计的主要目的是找到一个近似模型替代真实模型 $f(\cdot)$. 很容易发现正交设计和均匀设计之间存在着本质上的差异, 但这两类设计表都是建立在 U 型设计上的.

定义 1.6 一个 U 型设计是一个 $n \times s$ 的矩阵, 记为 $U(n; q_1 \times \cdots \times q_s)$, 第 j 列的元素属于 $\{1, \cdots, q_j\}$ 且这些元素在该列出现的次数相同. 当某些 q_j 相等时, $U(n; q_1 \times \cdots \times q_s)$ 可记为 $U(n; q_1^{r_1} \times \cdots \times q_m^{r_m})$, 这里 $r_1 + \cdots + r_m = s$. 当所有的q_j都相等且等于 q 时, U 型设计记为 $U(n; q^s)$, 此时对应的设计称为对称的, 否则称为非对称的或者是混水平 U 型设计. 记 $\mathcal{U}(n; q_1 \times \cdots \times q_s)$ 为所有 U 型设计 $U(n; q_1 \times \cdots \times q_s)$ 的集合. 类似地, 我们有符号 $\mathcal{U}(n; q_1^{r_1} \times \cdots \times q_m^{r_m})$ 和 $\mathcal{U}(n; q^s)$.

根据 U 型设计的定义, 很容易知道每一个 q_j 都能整除 n. 设 $\boldsymbol{U}=(u_{ij})$ 是 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 中的一个 U 型设计. 作如下变换

$$x_{ij}=\frac{u_{ij}-0.5}{q_j},\quad i=1,\cdots,n,\ j=1,\cdots,s. \tag{1.16}$$

记 $\boldsymbol{X}_u=(x_{ij})$, 称 $\boldsymbol{X}_u$ 为 $\boldsymbol{U}$ 的**诱导矩阵**. 矩阵 $\boldsymbol{X}_u$ 中的 n 行是 $[0,1]^s$ 上的 n 个点.

定义 1.5 中的第二个条件要求选择的试验点在某种意义上均匀地分布在试验域中. 为了满足这个条件, 需要一些均匀性度量. 文献中大部分均匀性度量是定义在 $[0,1]^s$ 上. 第 2 章将介绍各种偏差作为均匀性度量. 对于某偏差 D, 通过诱导矩阵的均匀性来定义 U 型设计的均匀性, 即

$$D(\boldsymbol{U})=D(\boldsymbol{X}_u). \tag{1.17}$$

定义 1.7 称设计 $\boldsymbol{U}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$ 为在 D 度量下的均匀设计, 假如

$$D(\boldsymbol{U})=\min_{\boldsymbol{V}\in\mathcal{U}(n;q_1\times\cdots\times q_s)}D(\boldsymbol{V}),$$

并记 $\boldsymbol{U}$ 为 $U_n(q_1\times\cdots\times q_s)$.

表 1.4 和表 1.5 分别给出了 $U_{12}(12^4)$(图 1.4) 和 $U_6(3^2\times 2)$ 的均匀设计表.

给定 (n,q,s), 对应的均匀设计并不唯一, 因为通过置换设计表的行和列, 每一类偏差是不变的.

表 1.4 $U_{12}(12^4)$

编号	1	2	3	4
1	1	10	4	7
2	2	5	11	3
3	3	1	7	9
4	4	6	1	5
5	5	11	10	11
6	6	9	8	1
7	7	4	5	12
8	8	2	3	2
9	9	7	12	8
10	10	12	6	4
11	11	8	2	10
12	12	3	9	6

表 1.5　$U_6(3^2 \times 2)$

编号	1	2	3
1	1	1	1
2	2	1	2
3	3	2	1
4	1	2	2
5	2	3	1
6	3	3	2

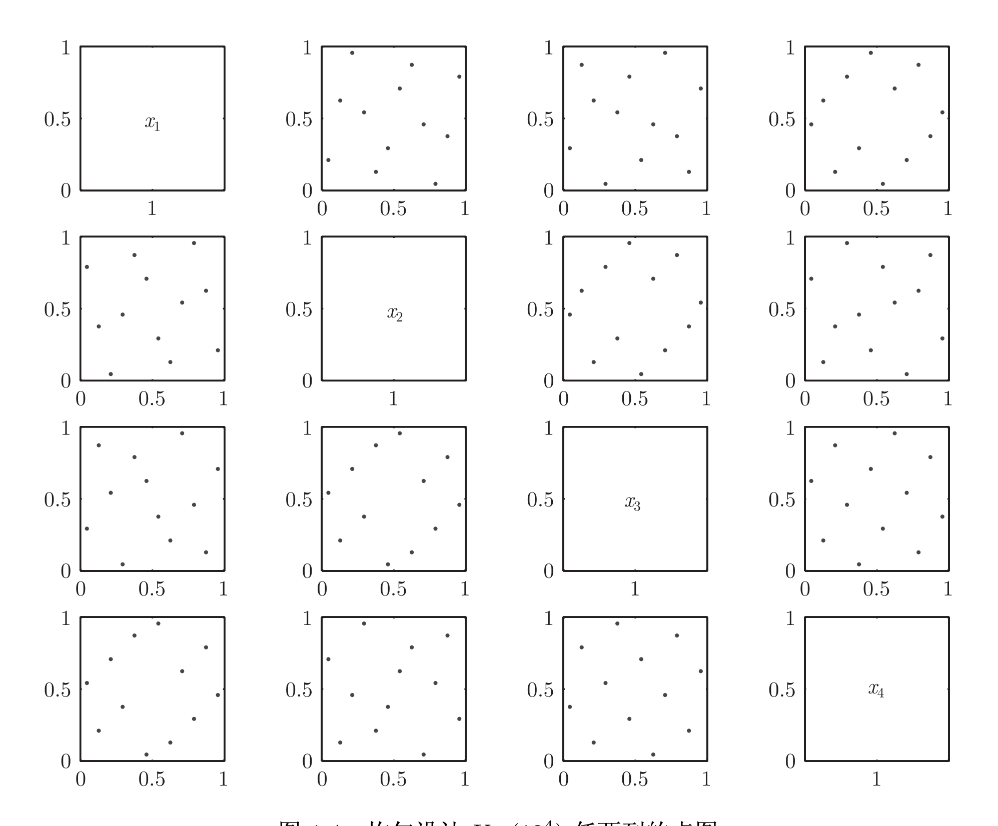


图 1.4　均匀设计 $U_{12}(12^4)$ 任两列的点图

定义 1.8　*两个 U 型设计称为是等价的, 假如一个设计通过置换行或置换列可以得到另一个设计.*

我们不区分等价的均匀设计. 给定一个均匀设计表 $U_n(q^s)$, 通过置换这个设计的行和列, 可以得到 $n!s!$ 个等价设计. 第 3 章和第 4 章将讨论均匀设计的构造, 仅需要研究等价均匀设计中的一个设计.

1.3.3 线性回归模型

假定试验者已知响应 y 和因子 $\boldsymbol{x}=(x_1,\cdots,x_s)$ 之间关系满足如下真实模型:

$$y(\boldsymbol{x})=\beta_1 g_1(\boldsymbol{x})+\cdots+\beta_m g_m(\boldsymbol{x})+\varepsilon, \tag{1.18}$$

其中 $\boldsymbol{x}=(x_1,\cdots,x_s)\in\mathcal{X}$, $\mathcal{X}$ 是 R^s 中的试验域, 函数 $g_1,\cdots,g_m$ 已知, $\beta_1,\cdots,\beta_m$ 是未知参数, ε 是随机误差并满足 $E(\varepsilon)=0$ 和 $\mathrm{Var}(\varepsilon)=\sigma^2$. Kiefer(1959) 首次研究在回归模型下的试验设计问题. 模型 (1.18) 是 β 的线性函数, 它包括了很多有用的模型, 例如, 简单线性模型

$$y=\beta_0+\beta_1 x_1+\cdots+\beta_s x_s+\varepsilon; \tag{1.19}$$

二次模型

$$y=\beta_0+\sum_{i=1}^{s}\beta_i x_i+\sum_{1\leqslant i\leqslant j\leqslant s}\beta_{ij}x_i x_j+\varepsilon \tag{1.20}$$

和中心二次模型

$$y=\beta_0+\sum_{i=1}^{s}\beta_i(x_i-\bar{x}_i)+\sum_{1\leqslant i\leqslant j\leqslant s}\beta_{ij}(x_i-\bar{x}_i)(x_j-\bar{x}_j)+\varepsilon, \tag{1.21}$$

其中 $\bar{x}_i$ 是 x_i 的样本均值. 注意到函数 g_j 关于 $\boldsymbol{x}$ 可以是非线性的, 例如, $\exp(-x_i/2)$, $\log(x_j)$, $1/(10+x_i x_j)$, 等等. 假定想用模型 (1.18) 来拟合数据集 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}\in\mathcal{X}$, 那么模型可表示为

$$\boldsymbol{y}=\boldsymbol{G}\boldsymbol{\beta}+\boldsymbol{\varepsilon}, \tag{1.22}$$

其中

$$\boldsymbol{G}=\begin{bmatrix} g_1(\boldsymbol{x}_1) & \cdots & g_m(\boldsymbol{x}_1)\\ \vdots & & \vdots\\ g_1(\boldsymbol{x}_n) & \cdots & g_m(\boldsymbol{x}_n)\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_1\\ \vdots\\ \beta_m\end{bmatrix},\quad \boldsymbol{\varepsilon}=\begin{bmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{bmatrix}, \tag{1.23}$$

$\varepsilon_1,\cdots,\varepsilon_n$ 为随机误差. 数据可表示为一个 $n\times s$ 的矩阵

$$\boldsymbol{X}=\begin{bmatrix} x_{11} & \cdots & x_{1s}\\ \vdots & & \vdots\\ x_{n1} & \cdots & x_{ns}\end{bmatrix}=\begin{bmatrix}\boldsymbol{x}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{x}_n^{\mathrm{T}}\end{bmatrix}.$$

称矩阵 $\boldsymbol{X}$ 为**设计矩阵**. 矩阵 $\boldsymbol{G}$综合了模型和数据的相关信息, 因此称之为**结构矩**

阵, 称 $\boldsymbol{M}=\boldsymbol{M}(\mathcal{P})=\dfrac{1}{n}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}$ 为信息矩阵. $\boldsymbol{\beta}$ 的最小二乘估计及其协方差矩阵分别为

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{y},\quad \mathrm{Cov}(\hat{\boldsymbol{\beta}})=\frac{\sigma^2}{n}\boldsymbol{M}^{-1}. \tag{1.24}$$

显然, 我们希望 $\mathrm{Cov}(\hat{\boldsymbol{\beta}})$ 在一定的意义下尽可能小. 这等价于要最大化 $\boldsymbol{M}(\mathcal{P})$. 由于 $\boldsymbol{M}$ 是一个 $m\times m$ 的矩阵, 比较矩阵的大小有许多方法, 其中一个比较简捷的方式就是寻找一个关于 $\boldsymbol{M}$ 的尺度函数作为准则, 记尺度函数为 $\phi(\boldsymbol{M}(\mathcal{P}))$, 以及寻找在设计空间上满足 $\phi(\boldsymbol{M})$ 最大化的 ϕ 最优设计. 文献中已有许多准则被用来寻找最优设计, 例如:

(1) D-最优性: 最大化 $\boldsymbol{M}$ 的行列式. 在多元分析中, 协方差的行列式称为广义方差. D-最优性等价于对任意的 $a^2>0$, 最小化置信椭球 $\{\boldsymbol{\beta}:(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{\mathrm{T}}\boldsymbol{M}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\leqslant a^2\}$ 的体积, 其中 $\hat{\boldsymbol{\beta}}=(\hat{\beta}_1,\cdots,\hat{\beta}_m)^{\mathrm{T}}$.

(2) A-最优性: 最小化 $\boldsymbol{M}^{-1}$ 的迹, 等价于最小化 $\hat{\beta}_1,\cdots,\hat{\beta}_m$ 的方差的和.

(3) E-最优性: 最小化 $\boldsymbol{M}^{-1}$ 的最大特征值或最小化 $\{\max_{||\boldsymbol{\alpha}||=1}\mathrm{Var}(\boldsymbol{\alpha}^{\mathrm{T}}\hat{\boldsymbol{\beta}})\}$.

(4) G-最优性: 最小化在试验域上预测响应的最大方差. 假设 $\hat{y}(\boldsymbol{x})$ 是在 $\boldsymbol{x}\in\mathcal{X}$ 处的预测响应, 即 $\hat{y}(\boldsymbol{x})=\sum_{j=1}^{m}\hat{\beta}_j g_j(\boldsymbol{x})$, 则 $\hat{y}(\boldsymbol{x})$ 的方差为

$$\mathrm{Var}(\hat{y}(\boldsymbol{x}))=\frac{\sigma^2}{n}\boldsymbol{g}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{g}(\boldsymbol{x}), \tag{1.25}$$

其中

$$\boldsymbol{g}(\boldsymbol{x})=(g_1(\boldsymbol{x}),\cdots,g_m(\boldsymbol{x}))^{\mathrm{T}}, \tag{1.26}$$

G-最优性就是最小化 $\{\max_{\boldsymbol{x}\in\mathcal{X}}\boldsymbol{g}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{g}(\boldsymbol{x})\}$.

矩阵的 “行列式”“迹” 和 “特征值” 等概念可在线性代数的教材中找到. 希望了解更多有关最优设计及其相关的理论的读者, 可查阅 (Atkinson and Donev, 1992) 与 (Pukelsheim, 1993).

若真实模型是已知的, 那么最优设计是最好的设计. 最优设计有很多非常好的性质, 但缺乏模型的稳健性. 当真实模型已知时, 要找到相应的最优设计往往会遇到一些困难. 当真实模型选错时, 最优设计可能表现不太好. 因此, 最优回归设计具有不稳健性.

由于 D-最优设计能估计出参数置信域的最小容量, 因此 D-最优设计可以改善估计的精度. 然而, 需要指出 D-最优设计存在如下问题. 假设所考虑的动力学模型具有非线性, 因此, 此时的 D-最优设计只是局部最优的. 换句话说, 它依赖于参数的先验选择. 假如参数初始值的估计值与真实值相差较远, 那么结果可能不太好. 但在实际中, 对于一个不熟悉的化学反应的动力学参数 (参考例 1.5), 我们往往知之

甚少, 因此, 选择合适的初始参数是 D-最优设计遇到的一个实际问题. 自然地, 我们想寻找其他试验方法使得对参数初始值不敏感, 同时也可以获得参数估计满意的精度. Xu 等 (2000) 对这个化学反应考虑了 D-最优设计 (DOD)、正交设计 (OD) 和均匀设计 (UD) 三种设计, 并比较了它们各自的表现. 例如, 一个化学反应的动力学模型假定是非线性模型, 为了刻画此模型, 每个因子取两个或三个水平看起来似乎太少了. 然而, UD 具有如下优势: 当你需要增加少量的试验次数时, UD 可以提供尽可能多的水平数来选择. 所以, UD 特别适合非线性模型. 因此, 在可逆反应中的动力学模型中, UD 有望对动力学模型的参数给出比较好的估计.

1.3.4 非参数回归模型

当试验者对于内在模型没有任何先验信息, 而又希望探索响应 y 和因子$(x_1, \cdots, x_s)$ 之间的关系时, 如下非参数回归模型可供考虑:

$$y(\boldsymbol{x}) = f(\boldsymbol{x}) + \varepsilon = f(x_1, \cdots, x_s) + \varepsilon, \tag{1.27}$$

其中函数 f 是未知的, 随机误差 ε 满足 $E(\varepsilon) = 0$ 和 $\mathrm{Var}(\varepsilon) = \sigma^2$, σ^2 是未知的. 我们希望在每个 $\boldsymbol{x}$ 处估计 $y(\boldsymbol{x})$. 一个自然的想法是在试验域上均匀地散布试验点. 这个思想产生了一些统计学术语: "空间填充设计", 包括 "均匀设计" 和 "拉丁超立方体抽样" (其定义可参见定义 2.1).

我们通过一个例子来说明对于同一个问题可以采用上述三类设计.

例 1.9(Weibull 生长模型) 在一个生物试验里, 我们希望探索生长时间 (x) 和响应 (y) 之间的联系. 假定内在模型

$$y = y(x) = 1 - e^{-2x^2}, \quad x \in [0, 2] \tag{1.28}$$

是未知的. 图 1.5 给出了此生长曲线图. 注意到随机误差将会影响响应 y. 对于该试验, 内在模型变成

$$y = y(x) = 1 - e^{-2x^2} + \varepsilon, \quad x \in [0, 2], \tag{1.29}$$

其中随机误差 $\varepsilon \sim N(0, \sigma^2)$, σ^2 是未知的. 图 1.6 给出了因子设计图, 其中因子水平数选定为四个水平, 在每个水平上重复三次试验. 通过该设计, 可以估计总体均值 μ 和 x 在这四个水平的主效应以及方差 σ^2. 如果在试验数据的分析中选择了如下的二次回归模型

$$y(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon, \tag{1.30}$$

那么相应的 D-最优设计如图 1.7 所示, 其中虚线是拟合曲线. 我们发现拟合效果不是很好, 模型 (1.30) 与真实模型偏离较远. 如果选择一个三次回归模型并应用相

关的 D-最优设计, 那么相应的结果会变得好些. 所以最优设计对模型的变化不具有稳健性. 一个 12 次试验的均匀设计如图 1.8 所示, 其中的虚线为由一个多项式回归模型得到的拟合曲线. 我们发现除了在 0 和 2 邻近的一些小区间外, 拟合效果很好. 由此可见, 在模型不确定的情况下, 试验的建模是一个非常重要且困难的问题.

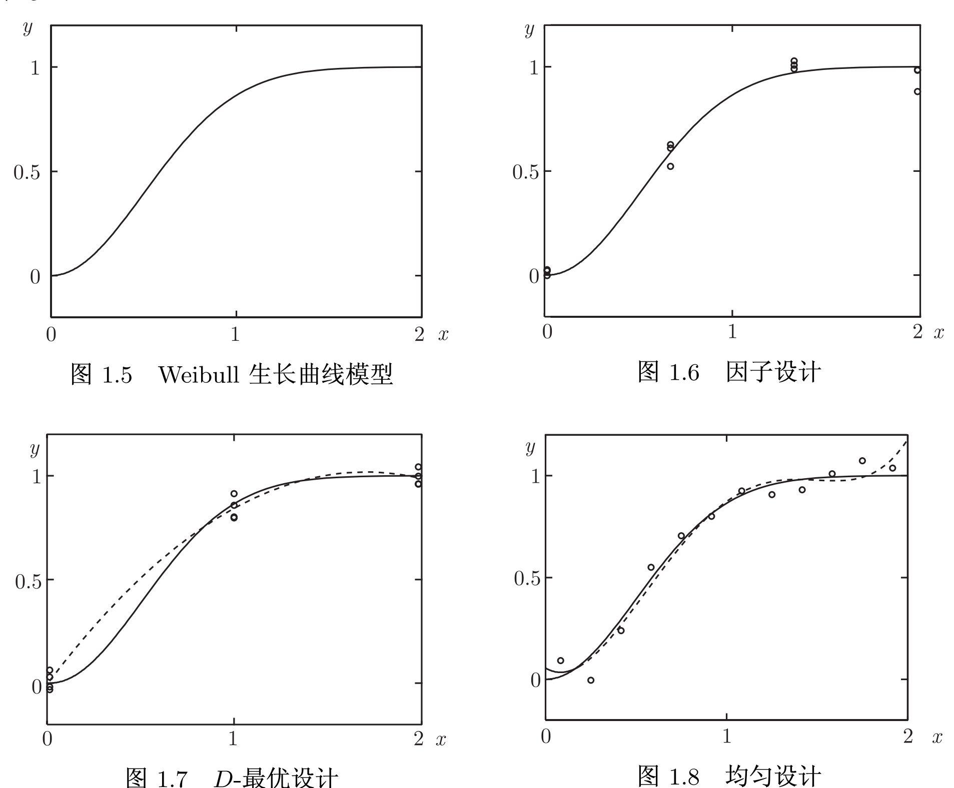


图 1.5　Weibull 生长曲线模型

图 1.6　因子设计

图 1.7　D-最优设计

图 1.8　均匀设计

1.3.5　回归模型的稳健性

在实际试验中, 由于所研究的问题的复杂性, 内在模型通常不完全已知. 用某种方式选择出来的设计, 如果它相对真实模型有小的偏差, 其试验结果变化不大, 那么称设计具有稳健性, 这是人们期待的. 也就是说, 内在模型或观测值小的变化应该只能引起推断统计过程小的变化. 常见的偏差可能来自于:

a) 假定模型是不正确的或不完全正确的;

b) 随机误差不服从假定分布 (例如, 正态分布);

c) 在试验域上随机误差的方差不是常数;

d) 存在不正常观测值 (异常值).

针对上述四种情形有小的偏离时, 一个好的试验设计应该具有稳健性. 一个设计具有这样的性质时称为*稳健设计*. 由于稳健性可能有不同的类别, 稳健设计依赖于具体的类别选项. Huber (1975), Box 和 Draper (1959), Marcus 和 Sacks (1978), Li 和 Notz (1982), Wiens (1990, 1991, 1992), Yue 和 Hickernell (1999) 与 Xie 和 Fang (2000) 对试验设计的稳健性和稳健设计进行了全面研究. 本书中仅局限于考虑模型的稳健性, 特别地, 我们只考虑下面的情形.

如果试验者知道内在模型接近于一个线性模型, 那么可考虑如下稳健的回归模型:

$$y(x) = \beta_1 f_1(x) + \cdots + \beta_m f_m(x) + h(x) + \varepsilon, \tag{1.31}$$

其中 $h(x)$ 表示模型 (1.18) 与真实模型的差别. 当函数 h 属于某个函数类时, 我们期待找到一个设计使得在特定情形下能获得 $\beta_1, \cdots, \beta_m$ 的最优估计. 例如, 我们定义一个损失函数, 可考虑最坏的情形和平均损失. 经证实, 均匀设计是针对模型变化的一种稳健设计.

在介绍均匀设计的理论之前, 我们将在 1.5 节中介绍一个模型不确定试验的真实案例.

1.4 字长型: 分辨度和最小低阶混杂

评价设计空间 $\mathcal{U}$ 里设计好坏的某个准则是一个关于 $\boldsymbol{U}$ 的函数, 记为 $\phi(\boldsymbol{U})$, 这里 $\boldsymbol{U} \in \mathcal{U}$. 每个准则都有自己的统计推断的特性或其他考虑. 有时候一个准则可能是含非负分量的向量 $(\phi_1(\boldsymbol{U}), \cdots, \phi_m(\boldsymbol{U}))$. 现在我们遇到的问题是如何在 $\mathbb{R}_+^m$ 上对两个非负向量 $\boldsymbol{x} = (x_1, \cdots, x_m)$ 与 $\boldsymbol{y} = (y_1, \cdots, y_m)$ 进行排序? 下面来介绍文献中常用的两种方法.

1.4.1 排序

(a) 字典序. 通过下面的方式序贯地比较 x_1 与 y_1, x_2 与 y_2, $\cdots$. 记 $\boldsymbol{x} \vdash \boldsymbol{y}$, 若 $x_1 < y_1$; 否则如果 $x_1 = y_1$ 且 $x_2 < y_2$; 否则如果 $x_i = y_i (i = 1, 2)$ 且 $x_3 < y_3$; $\cdots$; 否则如果 $x_i = y_i (i = 1, \cdots, m-1)$ 且 $x_m < y_m$. 更简洁的描述是: 记 $\boldsymbol{x} \vdash \boldsymbol{y}$, 若存在 $k, 1 \leqslant k \leqslant m$ 使得当 $i < k$ 时 $x_i = y_i$ 且 $x_k < y_k$; 记 $\boldsymbol{x} \models \boldsymbol{y}$ 若 $x_i \leqslant y_i, 1 \leqslant i \leqslant m$. 注意 "$\boldsymbol{x} \models \boldsymbol{y}$" 意味着 "$\boldsymbol{x} = \boldsymbol{y}$" 或 "$\boldsymbol{x} \vdash \boldsymbol{y}$". 这里, 关系 "$\models$" 和 "$\vdash$" 可分别看成是经典代数中的 "$\leqslant$" 和 "$<$".

(b) 优势序. 下面简单回顾*优势理论* (参见 (Marshall and Olkin, 1979)). 对任意

的正常数 c, 定义一个非负向量的集合

$$\mathcal{X}_c = \left\{ \boldsymbol{x} = (x_1, \cdots, x_m) : \boldsymbol{x} \in \mathbb{R}_+^m, \sum_{i=1}^m x_i = c \right\}.$$

对非负向量 $\boldsymbol{x} \in \mathcal{X}_c$, 记它的递增顺序统计量为 $x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(m)}$. 对任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_c$, 若

$$\sum_{r=1}^k x_{(r)} \geqslant \sum_{r=1}^k y_{(r)}, \quad k = 1, 2, \cdots, m, \tag{1.32}$$

我们说 $\boldsymbol{y}$ 优势于$\boldsymbol{x}$ 并记为 $\boldsymbol{x} \preceq \boldsymbol{y}$. 若存在某个 k 使得严格不等式 $\sum_{r=1}^k x_{(r)} > \sum_{r=1}^k y_{(r)}$ 成立, 则记为 $\boldsymbol{x} \prec \boldsymbol{y}$. 称 $\mathbb{R}_+^m$ 上的实值函数 Ψ 是 Schur凸的, 若对满足 $\boldsymbol{x} \preceq \boldsymbol{y}$ 的每一对 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^m$ 有 $\Psi(\boldsymbol{x}) \leqslant \Psi(\boldsymbol{y})$ 成立. 显然, $\Psi(\boldsymbol{x})$ 关于其分量是对称的, 即关于 $x_1, \cdots, x_m$ 的置换是不变的. 我们主要关心下面Schur凸函数的可分凸类:

$$\Psi(\boldsymbol{x}) = \sum_{r=1}^m \psi(x_r), \quad \psi \text{ 在 } \mathbb{R}_+ \text{ 上是凸的}, \tag{1.33}$$

其某个单调映射 $g(\Psi(\boldsymbol{x}))$ 在 $\mathbb{R}_+$ 上也是凸的. 下面的引理对获得一个准则的一些下界是非常重要的.

引理 1.1 假定向量 $\bar{\boldsymbol{x}} = (c/m, \cdots, c/m)$ 属于集合 $\mathcal{X}_c$ 且在优化顺序下是最小的, 即对任意的 $\boldsymbol{x} \in \mathcal{X}_c$ 有 $\bar{\boldsymbol{x}} \preceq \boldsymbol{x}$. 对任意的 Schur 凸函数 Ψ, 有 $\Psi(\bar{\boldsymbol{x}}) \leqslant \Psi(\boldsymbol{x})$, 且 $\Psi(\bar{\boldsymbol{x}})$ 是 Ψ 准则下的一个下界.

为了把上述引理应用到试验设计中去, 假设 $\boldsymbol{x}$ 是 $\mathcal{X}_c$ 中的整数向量. 若 c/m 是一个整数, 引理 1.1 中的下界是紧的, 否则是不可达的. 假设 θ 和 f 分别为 c/m 的整数和小数部分, 另定义

$$\widetilde{\boldsymbol{x}} \equiv (\underbrace{\theta, \cdots, \theta}_{m(1-f)}, \underbrace{\theta+1, \cdots, \theta+1}_{mf})^{\mathrm{T}}.$$

由 (Dey and Mukerjee, 1999) 中的引理 5.2.1, 可得如下的下界.

引理 1.2 在上述记号下有 $\bar{\boldsymbol{x}} \preceq \widetilde{\boldsymbol{x}} \preceq \boldsymbol{x}$. 对于 $\mathcal{X}_c$ 上任意的一个可分凸函数 $\sum_{r=1}^m \psi(x_r)$, 其有如下紧的下界:

$$m(1-f)\psi(\theta) + mf\psi(\theta+1). \tag{1.34}$$

当 c/m 为整数时, 上述下界变为 $m\psi(\theta)$.

Zhang 等 (2005) 利用上述引理给出了一些准则许多好的下界.

1.4.2 定义关系

对于部分因子设计文献已有很多准则来筛选比较好的设计, 可参见 (Dey and Mukerjee, 1999; Wu and Hamada, 2009). 一个正规部分因子设计是由其定义关系确定的. 先来看一个典型的例子. 考虑一个含四个因子的两水平试验, 因子记为 A, B, C 和 D. 利用表 1.6 的 7 列构成的正交设计表 $L_8(2^7)$, 两个水平记为 -1 和 1 (简记为 $-$ 和 $+$). 将因子 A, B, C 和 D 放在 1, 2, 4 和 7 列上. 在正交设计表 $L_8(2^7)$ 上添加一列 I, 其元素全为 1, 即添加在表 1.6 的第 1 列. 这个正交设计表有下列特征:

1) 除去列 I, 每列都有相等的正号和负号;

2) 任意两列的点积是零;

3) 列 I 点乘以任意列保持该列不变;

4) 任意两列的点积可产生表中的某一列;

例如, $12 = 3, 57 = 2, 34 = 7$, 这里 12 表示点积 $1 \cdot 2$, 其他的类似.

5) 1, 2, 4 和 7 列的点积为列 I, 并记为 $I = 1247$. 若把 A, B, C 和 D 分别放在列 1, 2, 3, 4 上, 也可表示为 $I = ABCD$. 这里 $ABCD$ 称为这个特定部分的生成元, $I = ABCD$ 称为*定义关系*.

6) 在 $I = ABCD$ 两边同时点乘 A 有 $A = BCD$. 类似地有

$$A = BCD, \quad B = ACD, \quad C = ABD, \quad D = ABC, \tag{1.35}$$

$$AB = CD, \quad AC = BD, \quad AD = BC, \tag{1.36}$$

这意味着主效应 A 与 交互效应 BCD 是混杂的, 交互效应 AB 与 CD 混杂, 等等. 所有的混杂关系都可从定义关系 $I = ABCD$ 中获得. 若所有的三阶交互效应可忽略, 则混杂关系变为

$$AB = CD, \quad AC = BD, \quad AD = BC.$$

此时, 所有的主效应是可估的, 所有的二阶交互效应相互混杂.

表 1.6 正交设计 $L_8(2^7)$

编号	I	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1	1
2	1	1	1	1	−1	−1	−1	−1
3	1	1	−1	−1	1	1	−1	−1
4	1	1	−1	−1	−1	−1	1	1
5	1	−1	1	−1	1	−1	1	−1
6	1	−1	1	−1	−1	1	−1	1
7	1	−1	−1	1	1	−1	−1	1
8	1	−1	−1	1	−1	1	1	−1

1.4.3 字长型和分辨度

许多设计是由多个定义关系来确定的. 下面来看一个例子.

例 1.10 考虑一个 2^{5-2} 的 FFD, 如表 1.7 所示. 2^{s-k} 的 FFD 是一个含 $n=2^{s-k}$ 次试验和 k 个两水平因子的部分因子设计. 根据点乘的规则, 我们发现该设计存在如下关系

$$4=12,\quad 5=13. \tag{1.37}$$

这些关系意味着

$$I=124=135=2345, \tag{1.38}$$

即为该设计的定义关系. 关系 $I=2345$ 可由关系 $I=124$ 和 $I=135$ 获得. 由这些定义关系, 所有的混杂关系可表示如下

$$\begin{aligned}1&=24=35=12345,\\2&=14=345=1235,\\3&=15=245=1234,\\4&=12=235=1345,\\5&=13=234=1245.\end{aligned}$$

表 1.7 2^{5-2} FFD

编号	1	2	3	4	5
1	1	1	1	1	1
2	1	1	−1	1	−1
3	1	−1	1	−1	1
4	1	−1	−1	−1	−1
5	−1	1	1	1	−1
6	−1	1	−1	1	1
7	−1	−1	1	−1	−1
8	−1	−1	−1	−1	1

假设 $\boldsymbol{U}$ 是由 k 个定义关系或定义字确定的一个 2^{s-k} 的 FFD, 其中字由表示因子的 $1,2,\cdots,s$ 或 $A,B,\cdots$ 的字母组成. 一个字中字母的个数为该字的字长, k 个定义字构成的群称为定义对照子群. 该群有 2^k-1 个字和一个恒等元 I.

令 $A_i(\boldsymbol{U})$ 为设计 $\boldsymbol{U}$ 的定义对照子群中字长为 i 的字的个数, 称向量

$$W(\boldsymbol{U})=(A_1(\boldsymbol{U}),\cdots,A_s(\boldsymbol{U})) \tag{1.39}$$

为 $\boldsymbol{U}$ 的字长型. 设计 $\boldsymbol{U}$ 的字长型显示了其统计推断能力. $\boldsymbol{U}$ 的分辨度由 Box 和 Hunter (1961a, 1961b) 提出, 其定义为具有正数 $A_t(\boldsymbol{U})$ 的最小的 t. 分辨度为 III 和 IV 的设计在实际中应用比较广泛.

分辨度为 III 的设计 没有主效应与其他主效应混杂, 但主效应与二因子交互效应混杂, 二因子交互效应之间可能相互混杂.

分辨度为 IV 的设计 没有主效应与其他主效应或二因子交互效应混杂, 但二因子交互效应之间可能相互混杂.

例 1.11 *一个定义关系为 $I = 123$ 的 2^{3-1} 设计, 其分辨度为 III, 其字长型为 $W = (0, 0, 1)$; 一个定义关系为 $I = 1234$ 的 2^{4-1} 设计, 其分辨度为 IV, 其字长型为 $W = (0, 0, 0, 1)$; 例 1.10 中的设计的分辨度为 III, 字长型为 $W = (0, 0, 2, 1, 0)$.*

例 1.12 *考虑两个 2^{5-1} 的正交设计, 记为 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$. 它们的定义关系分别为*

$$\boldsymbol{U}_1 : I = 4567 = 12346 = 12357,$$
$$\boldsymbol{U}_2 : I = 1236 = 1457 = 234567.$$

它们有相同的分辨度 IV, 但它们有不同的字长型:

$$W(\boldsymbol{U}_1) = (0, 0, 0, 1, 2, 0, 0), \quad W(\boldsymbol{U}_2) = (0, 0, 0, 2, 0, 1, 0).$$

因此, $\boldsymbol{U}_1$ 比 $\boldsymbol{U}_2$ 有较小的混杂. 事实上, 利用 $\boldsymbol{U}_1$ 我们发现有三对二因子交互效应混杂, 即 $45 = 67, 46 = 57, 47 = 56$, 而利用 $\boldsymbol{U}_2$ 我们发现有六对二因子交互效应混杂, 即 $12 = 36, 13 = 26, 16 = 23, 14 = 57, 15 = 47, 17 = 45$. 因此, 我们需要更数学化的准则来筛选好的设计, 其中最小低阶混杂准则就是最为常见的准则之一.

1.4.4 最小低阶混杂准则及其推广

*最小低阶混杂准则*由 Fries 和 Hunter (1980) 提出, 现已被广泛应用于正规部分因子设计. 最小低阶混杂准则也是基于字长型来定义的. 因为我们接下来会给出最小低阶混杂准则的一般形式, 所以这里就省略其定义. 最小低阶混杂准则假设效应遗传原则成立, 它趋向于从低维到高维序贯地最小化不可忽略的交互效应的混杂. 许多作者已经在寻找最小低阶混杂设计方面做了深入研究, 例如, (Cheng and Mukerjee, 1998; Fang and Mukerjee, 2000; Mukerjee and Wu, 2001; Cheng and Tang, 2005).

最近, 非正规设计的广义混杂的定义已成为一个研究热点. Tang 和 Deng (1999) 对于非正规两水平设计提出了最小 G_2 混杂准则, Xu 和 Wu (2001) 与 Ma 和 Fang (2001) 分别独立地提出了广义字长型及其相关的准则. 这方面重要的工作还包括 (Deng and Tang, 2002; Cheng et al., 2002; Ye, 2003; Fang and Zhang, 2004).

根据 (Xu and Wu, 2001), 考虑如下 ANOVA 模型:

$$\boldsymbol{y} = \alpha_0 \mathbf{1}_n + \boldsymbol{X}_{(1)}\boldsymbol{\alpha}_1 + \cdots + \boldsymbol{X}_{(s)}\boldsymbol{\alpha}_s + \boldsymbol{\varepsilon},$$

其中 $\boldsymbol{y}$ 是 n 个观测值向量, $\mathbf{1}_n$ 是元素为 1 的 n 维向量, α_0 是截距项, $\boldsymbol{\alpha}_j$ 是所有 j 因子交互效应的向量, $\boldsymbol{X}_{(j)}$ 是 $\boldsymbol{\alpha}_j$ 的对照系数矩阵, $\boldsymbol{\varepsilon}$ 是独立的随机误差向量. 对于一个 $U(n; q_1 \times \cdots \times q_s)$ 设计 $\mathcal{P}$, 假设 $\boldsymbol{X}_{(j)} = (x_{ik}^j)$ 为所有 j 因子对照系数组成的矩阵, $j = 0, \cdots, s$.

定义 1.9　*定义*

$$A_j(\mathcal{P}) = n^{-2} \sum_k \left| \sum_{i=1}^n x_{ik}^j \right|^2 . \tag{1.40}$$

称向量 $W(\mathcal{P}) = (A_1(\mathcal{P}), \cdots, A_s(\mathcal{P}))$ 为 $\mathcal{P}$ 的广义字长型.

该定义源于 (Xu and Wu, 2001). $\mathcal{P}$ 的*分辨度*是 $W(\mathcal{P})$ 中具有正数 $A_j(\mathcal{P})$ 的最小 j. 假设 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 为两个不同的设计, t 是使得 $A_t(\boldsymbol{U}_1) \neq A_t(\boldsymbol{U}_2)$ 的最小整数. 称 $\boldsymbol{U}_1$ 比 $\boldsymbol{U}_2$ *有更小的广义混杂*, 若 $A_t(\boldsymbol{U}_1) < A_t(\boldsymbol{U}_2)$. 若在设计空间中没有其他设计比 $\boldsymbol{U}$ 有更小的广义混杂, 称设计 $\boldsymbol{U}$ 有*广义最小低阶混杂*(GMA).

注意到, 对于一个 U 型设计 $\boldsymbol{U} \in \mathcal{U}(n; q^s)$, 基于编码理论, Ma 和 Fang (2001) 提出了*广义字长型*的另一种定义, 记为 $W^g(\boldsymbol{U}) = \{A_1^g(\boldsymbol{U}), \cdots, A_s^g(\boldsymbol{U})\}$ 及相应的*广义最小低阶混杂准则*. 由于对于一个 U 型设计 $\boldsymbol{U} \in \mathcal{U}(n; q^s)$, 有

$$A_j^g(\boldsymbol{U}) = A_j(\boldsymbol{U})/(q-1), \quad j = 1, \cdots, s.$$

因此, 我们省略了该准则的细节.

对于广义字长型, 有如下事实:

(1) 对任意正规设计 $\boldsymbol{U}$, $A_j(\boldsymbol{U})$ 就变成了 (1.39) 中的 $A_j(\boldsymbol{U})$;

(2) 对任意的设计 $\boldsymbol{U}$ 有 $A_j(\boldsymbol{U}) \geqslant 0, 1 \leqslant j \leqslant s$;

(3) 对任意的 U 型设计, $A_1(\boldsymbol{U}) = 0$;

(4) 若 $\boldsymbol{U}$ 是一个强度为 t 的正交设计表, 则有 $A_j(\boldsymbol{U}) = 0$, $j = 1, \cdots, t$.

1.5　多因子试验中均匀设计的实施

在本节中, 我们将给出在一个未知模型的实体试验中实施均匀设计的具体过程, 并提供一个利用均匀设计进行试验的实例. 本节所介绍的试验是由 Zeng (1994) 具体实施的, 随后 Fang (2002) 利用这个实例来说明均匀设计的目的. 为了方便读者阅读, 在本节中重新组织了该试验的相关内容.

通过下面的步骤给出了多因子工业或实验室试验的一般指南:

步骤1 选择因子和试验域, 确定每个因子合适的水平数.

步骤2 选择一个合适均匀设计表, 使之能满足试验所需的因子个数和相应的水平数. 这很容易通过访问均匀设计网站或利用计算机软件实现.

步骤3 根据均匀设计表, 随机确定试验顺序并实施试验.

步骤4 利用一个或几个模型来拟合试验数据集. 回归分析、神经网络、小波、多元样条和经验 Kriging 模型等对建模是非常有用的.

步骤5 从所建立的模型中发现知识. 例如, 我们想找到因子水平的 "最佳" 组合, 使得在该组合下响应能最大化或最小化, 而且可以通过进一步的试验得到验证.

步骤6 深入研究: 通过数据分析获得的最新信息, 可能需要追加一些试验.

下面通过一个化工试验例子来描述该过程的步骤.

例 1.13 为了寻找到一个最佳水平组合以便能提高化合物的产量, 我们需要安排、进行一个化学试验.

本试验需要考虑如下四个因子: 甲醛量 (x_1)、反应温度 (x_2)、反应时间 (x_3) 和碳酸钾量(x_4). 响应变量为产量(y). 试验域取为 $\mathcal{X}=[1.0,5.4]\times[5,60]\times[1.0,6.5]\times[15,70]$, 每个因子在该试验中取 12 个水平. 四个因子和它们的水平呈现在表 1.8 中.

表 1.8 因子和水平

因子	单位	水平
甲醛量 (x_1)	摩尔	1.0, 1.4, 1.8, 2.2, 2.6, 3.0, 3.4, 3.8, 4.2, 4.6, 5.0, 5.4
反应温度 (x_2)	摄氏度	5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60
反应时间 (x_3)	小时	1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5
碳酸钾量 (x_4)	毫升	15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70

选择一个形如 $U_n(12^4)$ 的均匀设计表, 并将每个因子分配到表的列中, 使得四个因子分到不同的列中. 每个因子的 12 个水平记为 $1,2,\cdots,12$, 并转换为对应的实际水平. 这样的设计列在表 1.9. 为简便起见, 我们也暂时称表 1.9 为表. 随机化 12 个水平组合并把试验顺序列在表的第 1 列, 根据表中的水平组合进行 12 次试验并将相应的产量 y 记录在表的最后一列.

针对各种不同的试验目的, 选择一个好的初始模型将会非常有用, 特别是在寻找使得产量最大的四个因子最好水平组合, 这是该试验数据分析的主要目标. 在 12 个响应中最好的结果是 $y_1 = 18.36\%$, 其水平组合为 $x_1 = 5.0, x_2 = 40$, $x_3 = 1.5$, $x_4 = 60$. 这可作为一个基准. 注意到这 12 次试验代表该试验域中的 $12^4 = 20736$ 个全体水平组合. 很有可能找到另一个水平组合的产量高于 $y_1 = 18.36\%$. 为此, 若找到一个初始模型来近似真实模型, 那么就有可能找到一个产量更高的水平组合.

表 1.9 $U_{12}(12^4)$ 和相关的设计

试验号	x_1	x_2	x_3	x_4	y
5	1.0 (1)	50 (10)	2.5 (4)	45 (7)	0.0795
6	1.4 (2)	25 (5)	6.0 (11)	25 (3)	0.0118
10	1.8 (3)	5 (1)	4.0 (7)	55 (9)	0.0109
7	2.2 (4)	30 (6)	1.0 (1)	35 (5)	0.0991
11	2.6 (5)	55 (11)	5.5 (10)	65 (11)	0.1266
9	3.0 (6)	45 (9)	4.5 (8)	15 (1)	0.0717
8	3.4 (7)	20 (4)	3.0 (5)	70 (12)	0.1319
3	3.8 (8)	10 (2)	2.0 (3)	20 (2)	0.0900
2	4.2 (9)	35 (7)	6.5 (12)	50 (8)	0.1739
4	4.6 (10)	60 (12)	3.5 (6)	30 (4)	0.1176
1	5.0 (11)	40 (8)	1.5 (2)	60 (10)	0.1836*
12	5.4 (12)	15 (3)	5.0 (9)	40 (6)	0.1424

由于我们对内在模型缺乏了解, 因此可以考虑用非参数回归模型 (1.27) 来建模. 已有许多方法来寻找好的初始模型, 读者可以通过查阅 (Fang et al., 2006a) 来进行深入细致的了解. 对于含有多个因子的试验, 通常先考虑一组基函数 $\{B_0(\boldsymbol{x}), B_1(\boldsymbol{x}), \cdots\}$ 和基于基函数的最大模型

$$\hat{g}(\boldsymbol{x}) = \beta_1 B_1(\boldsymbol{x}) + \beta_2 B_2(\boldsymbol{x}) + \cdots + \varepsilon. \tag{1.41}$$

然后利用变量选择技术, 将 (1.41) 的一个子模型作为初始模型. 在本例中, 仅考虑线性和二次回归模型这两种情形. 由于线性回归模型形式简单, 因此首先尝试用如下线性回归模型来拟合数据:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4.$$

基于表 1.9 中的数据, 可以得到

$$\hat{y} = -0.0533 + 0.0281x_1 + 0.0010x_2 - 0.0035x_3 + 0.0011x_4, \tag{1.42}$$

其方差分析表见表 1.10. 从方差分析表中我们发现: 模型 (1.42) 包含一个不显著项 “x_3”, 其 p 值为 0.4962. 因此, 我们需要把该项从模型中去掉. 通过回归分析中向后消除技术 (Miller, 2002), 可以得到如下模型

$$\hat{y} = 0.0107 + 0.0289x_1,$$

表1.10 模型(1.42)的方差分析表

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
Model	4	0.0274	0.0069	8.2973	0.0086
Error	7	0.0058	0.0008		
C Total	11	0.0332			

Type III Tests					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
X1	1	0.0180	0.0180	21.8021	0.0023
X2	1	0.0033	0.0033	3.9496	0.0872
X3	1	0.0004	0.0004	0.5150	0.4962
X4	1	0.0046	0.0046	5.6248	0.0495

其中 $R^2 = 57.68\%$ 和 $s^2 = 0.0014$. 该模型与试验者的经验是不一致的, 因为有三个因子没有出现在模型中. 进而考虑如下更灵活的二次回归模型作为另一个最大模型:

$$E(y) = \beta_0 + \sum_{i=1}^{4} \beta_i x_i + \sum_{i \leqslant j} \beta_{ij} x_i x_j. \tag{1.43}$$

注意到该模型的未知参数个数大于试验次数, 因此模型是不可估的. 然而, 该模型可以为我们进行数据分析提供了基础, 其某些子模型可能会很好地拟合试验数据. 接下来将要寻找 (1.43) 的一个子模型来作为初始模式. 利用变量选择中的 MAXR 技术, 找到如下好的子模型:

$$\begin{aligned}\hat{y} = &\ 0.0446 + 0.0029x_2 - 0.0260x_3 + 0.0071x_1x_3 \\ &+ 0.000036x_2x_4 - 0.000054x_2^2,\end{aligned} \tag{1.44}$$

其中 $R^2 = 97.43\%$ 和 $s^2 = 0.0001$, 相应的方差分析表呈现在表 1.11 中.

统计诊断对于初始模型的检验是非常有用的. 通常情况下回归模型会满足某些特定的假设 (如正态性、常数方差等等), 但有的时候这些假设是值得怀疑的. 因此, 需要对这些假设进行统计诊断. 例如, 模型是否正确? 是否有异常值? 方差是否为常数? 误差是否服从正态分布? 如果某些假设失效, 我们必须进行修正. 读者可查阅 (Mayers, 1990) 或 (Cook, 1986) 来了解统计诊断的一些基本知识. 通过统计诊断我们得知模型 (1.44) 是可以接受的. 其具体细节在此省略.

在文献中, 如下的中心二次回归模型

$$E(y) = \beta_0 + \sum_{i=1}^{4} \beta_i (x_i - \bar{x}_i) + \sum_{i \leqslant j} \beta_{ij} (x_i - \bar{x}_i)(x_j - \bar{x}_j) \tag{1.45}$$

表1.11　模型(1.44)的方差分析表

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
Model	5	0.0323	0.0065	45.5461	0.0001
Error	6	0.0009	0.0001		
C Total	11	0.0332			

Type III Tests					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
X2	1	0.0014	0.0014	10.1949	0.0188
X3	1	0.0125	0.0125	88.0883	0.0001
X1X3	1	0.0193	0.0193	135.5636	0.0001
X2X2	1	0.0024	0.0024	16.8276	0.0063
X2X4	1	0.0062	0.0062	43.6923	0.0006

也被考虑为感兴趣的最大模型, 其中 $\bar{x}_i$ 是 x_i 的样本均值. 本例中, $\bar{x}_1 = 3.2, \bar{x}_2 = 32.5, \bar{x}_3 = 3.75$ 和 $\bar{x}_4 = 42.5$. 再次利用模型选择技术, 最终的模型推荐为

$$\begin{aligned}\hat{y} = {} & 0.1277 + 0.0281(x_1 - 3.2) + 0.000937(x_2 - 32.5) + 0.00114(x_4 - 42.5) \\ & + 0.00058(x_3 - 3.75)(x_4 - 42.5) - 0.000082(x_2 - 32.5)^2, \end{aligned} \tag{1.46}$$

其中 $R^2 = 97.05\%$ 和 $s^2 = 0.0002$, 相应的方差分析表呈现在表 1.12 中.

表1.12　模型(1.46)的方差分析表

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
Model	5	0.0322	0.0064	39.5466	0.0002
Error	6	0.0010	0.0002		
C Total	11	0.0332			

Type III Tests					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
X1	1	0.0180	0.0180	110.7654	0.0001
X2	1	0.0031	0.0031	19.2816	0.0046
X4	1	0.0046	0.0046	28.5141	0.0018
X3X4	1	0.0023	0.0023	13.8349	0.0099
X2X2	1	0.0047	0.0047	29.0842	0.0017

初始模型 (1.44) 和 (1.46) 可用来预测试验域中任意一点的响应, 也可以用来寻找最优的因子组合. 在*步骤*1 中的范围 $\mathcal{X}$ 内, 分别在模型 (1.44) 和 (1.46) 下来最大化 y, 即寻找 $x_i^*, i=1,\cdots,4$, 使得

$$\hat{y}(x_1^*,x_2^*,x_3^*,x_4^*)=\max_{\mathcal{X}}\hat{y}(x_1,x_2,x_3,x_4),$$

其中 $\hat{y}(x_1,x_2,x_3,x_4)$ 由 (1.44) 或 (1.46) 给出. 利用优化算法, 容易发现在模型 (1.44) 下, 当 $x_1^*=5.4, x_2^*=50.2, x_3^*=1$ 和 $x_4^*=70$ 时, 相应的响应 $\hat{y}(5.4,50.2,1,70)=19.3\%$达到最大; 在模型 (1.46) 下, 当 $x_1^*=5.4, x_2^*=43.9, x_3^*=6.5$ 和 $x_4^*=70$ 时, 相应的响应 $\hat{y}(5.4,43.9,6.5,70)=26.5\%$达到最大. 从上述分析来看, 模型 (1.46) 似乎要好一些, 但究竟哪个初始模型要好一些还需要进一步的试验来判断.

注意到两个最优点 $\boldsymbol{x}_1^*=(5.4,50.2,1,70)$ 和 $\boldsymbol{x}_2^*=(5.4,43.9,6.5,70)$ 都没有出现在试验方案 (表 1.9) 中, 因此需要追加一些试验来进一步验证我们的分析结果. 一种简单的办法就是在这两个最优点 $\boldsymbol{x}_1^*$ 和 $\boldsymbol{x}_2^*$ 再进行 m 次试验并比较它们相应的响应均值. 在这个试验中, 试验者在 $\boldsymbol{x}_1^*$ 和 $\boldsymbol{x}_2^*$ 又进行了三次试验, 并发现: 在 $\boldsymbol{x}_1^*$ 处, y 的平均值为 20.1%; 在 $\boldsymbol{x}_2^*$ 处, y 的平均值为 26.3%. 因此, 我们选择中心二次回归模型 (1.46) 作为最终的模型.

注意到初始模型都推荐 $x_1=5.4$ 和 $x_4=70$. 这意味着需要考虑增加 x_1 和 x_4 的试验水平上界. 试验者应考虑更进一步的研究并安排相应的试验.

1.6 均匀设计的应用

均匀试验设计 (简称均匀设计) 是一种使得试验点均匀地分布在试验域里的设计. 均匀设计由方开泰教授和王元院士于 1980 年提出来 (Fang, 1980; Wang and Fang, 1981), 现已被广泛应用于各种模型未知的实体和计算机试验中. 从文献中看到均匀设计已被广泛地用作:

1) 模型未知的部分因子设计;

2) 用于计算机试验的空间填充设计;

3) 给定模型的稳健设计;

4) 混料试验设计;

5) 超饱和设计.

*超饱和设计*是主效应和交互效应数大于试验次数的因子设计. 这类设计常称为筛选设计. 许多作者欣赏均匀设计的优点就在于均匀设计:

1) 用户有更多的设计选择;

2) 有现成的设计表格;

3) 可同时适用于因子试验和计算机试验;

4) 可接受内在模型信息较少的试验.

均匀设计与其他各种设计存在着联系, 如部分因子设计, 包括正交设计表、超饱和设计、稳健设计、组合设计和编码理论. 许多均匀设计表可在下面的网站上找到

http://www.math.hkbu.edu.hk/UniformDesign

或

http://web.stat.nankai.edu.cn/cms-ud/

习 题

1.1 比较实体试验和计算机试验, 并列举它们之间的差别.

1.2 什么是计算机试验的初始模型? 请给出初始模型的一些要求.

1.3 一般在一个试验中有许多变量. 请给出变量与试验设计中的因子之间的差别. 举例说明定量因子和定性因子. 环境变量和噪声变量之间的差别是什么?

1.4 考虑单因子试验的模型及其统计模型 (1.9)

$$y_{ij} = \mu_j + \varepsilon_{ij}, \quad j = 1, \cdots, q,\ i = 1, \cdots, n_j,$$

其中 ε_{ij} 是独立同分布 $N(0, \sigma^2)$. 设 $n = n_1 + \cdots + n_q$ 是试验数. 模型可以表示为一个线性模型 $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$.

1) 请详细给出 $\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\beta}$.

2) 给出 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 和 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$.

3) 给出 y_{ij} 和 $\boldsymbol{y}$ 的分布.

1.5 前面一个练习的模型可表示为 (1.10), 即

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij}, \quad j = 1, \cdots, q,\ i = 1, \cdots, n_j.$$

1) 把该模型改写成线性模型 $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ 并写出 $\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\beta}$.

2) 设 $\mathrm{SS}_E = \sum_{i=1}^{q}\sum_{j=1}^{r}(y_{ij} - \bar{y}_i)^2$ 为误差平方和, 其中 $n_1 = \cdots = n_q = r$. 证明 $E[\mathrm{SS}_E] = q(r-1)\sigma^2$.

1.6 对于二因子试验的 ANOVA 模型 (1.11), 即

$$\begin{aligned} y_{ijk} = &\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, \\ &i = 1, \cdots, q_1,\ j = 1, \cdots, q_2,\ k = 1, \cdots, r, \end{aligned}$$

其约束为

$$\begin{aligned} \alpha_1 + \cdots + \alpha_{q_1} = 0; \quad \beta_1 + \cdots + \beta_{q_2} = 0; \\ \sum_{i=1}^{q_1}(\alpha\beta)_{ij} = 0,\ j = 1, \cdots, q_2; \quad \sum_{j=1}^{q_2}(\alpha\beta)_{ij} = 0,\ i = 1, \cdots, q_1. \end{aligned}$$

证明:

1) 当 $q_1=q_2=2$ 时, 在 $\{(\alpha\beta)_{11},(\alpha\beta)_{12},(\alpha\beta)_{21},(\alpha\beta)_{22}\}$ 中只有一个线性独立交互效应.

2) 在 $\{(\alpha\beta)_{ij},i=1,2;j=1,\cdots,q_i\}$ 中有 $(q_1-1)(q_2-2)$ 个线性独立交互效应.

1.7 考虑三因子试验下面的模型

$$y_{ijk}=\mu+\alpha_i+\beta_j+\gamma_k+(\beta\gamma)_{jk}+\varepsilon_{ijk},\quad i=1,2,3,\ j=1,2,3,\ k=1,2,$$

其中 A,B,C 为因子,

y_{ijk}: 在 $A=A_i,B=B_j$ 和 $C=C_k$ 上的响应,

μ : 总体均值,

α_i : 因子 A 在水平 A_i 上的主效应,

β_j : 因子 B 在水平 B_j 上的主效应,

γ_k : 因子 C 在水平 C_k 上的主效应,

$(\beta\gamma)_{jk}$: 因子 B 和 C 在水平 B_j 与 C_k 上的交互效应,

ε_{ijk} : 在 $A=A_i$, $B=B_j$ 与 $C=C_k$ 上试验的随机误差, 且 $\varepsilon_{ijk}\sim N(0,\sigma^2)$.

请回答下列问题:

a) 给出 $\alpha_i,\beta_j,\gamma_k$ 与 $(\beta\gamma)_{jk}$ 的约束. 在 $\alpha_i,\beta_j,\gamma_k$ 与 $(\beta\gamma)_{jk}$ 中有多少个独立参数?

b) 改写模型形如 $\boldsymbol{y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 并给出 $\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta}$ 与 $\boldsymbol{\varepsilon}$, 其中 $\boldsymbol{\beta}$ 是问题 a) 中讨论的独立参数.

c) 给出平方和 $\mathrm{SS}_A,\mathrm{SS}_B,\mathrm{SS}_C,\mathrm{SS}_{B\times C},\mathrm{SS}_E$ 与 SS_T 的自由度.

d) 给出 $\mathrm{SS}_A,\mathrm{SS}_B,\mathrm{SS}_C,\mathrm{SS}_{B\times C},\mathrm{SS}_E$ 与 SS_T 的表达式.

1.8 正交性的概念出现在不同的领域中, 请回答下列问题:

1) 给出 R^d 中两个线段正交的定义.

2) 给出两个平面正交的定义.

3) 给出两个线性空间正交的定义.

4) 设 $\boldsymbol{X}$ 为正交设计表 $L_9(3^4)$ 的设计矩阵, 记 $\mathcal{L}_i(i=1,2,3,4)$ 为 $\boldsymbol{X}$ 的第 i 列生成的线性子空间, $\mathcal{L}_{ij}(1\leqslant i<j\leqslant 4)$ 为 $\boldsymbol{X}$ 的第 i 列与第 j 列生成的线性子空间. 证明: 若 $\mathcal{L}_{12}$ 与 $\mathcal{L}_{34}$ 是正交的, 则对于 $i\neq j$, $\mathcal{L}_i$ 与 $\mathcal{L}_j$ 是正交的.

1.9 给出下列设计的字长型模式和分辨度:

(1) 以 $I=ABCE=BCDF$ 为定义关系的 2^{6-2} 设计.

(2) 以 $I=ABCDF=ABDEG$ 为定义关系的 2^{7-2} 设计.

(3) 以 $I=ABCE=BCDF=ACDG$ 为定义关系的 2^{7-3} 设计.

1.10 请回答下列问题:

(1) 在 MATLAB 中有一个命令 “hadamard” 来生成一个 n 阶 Hadamard 矩阵. 利用这个命令来找到一个 8 阶 Hadamard 矩阵, 该矩阵可获得 $L_8(2^7)$.

(2) 若 $\boldsymbol{H}$ 是一个 Hadamard 矩阵, 则 $-\boldsymbol{H}$ 也是 Hadamard 矩阵.

(3) 若 $\boldsymbol{H}$ 是一个 Hadamard 矩阵, 设

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{H} & -\boldsymbol{H}\\ \boldsymbol{H} & \boldsymbol{H}\end{bmatrix}.$$

证明 $\boldsymbol{V}$ 是一个 Hadamard 矩阵.

1.11 计算下面设计 $L_9(3^4)$ 的任意两两不同行的 Hamming 距离, 请给出你的发现和推测.

编号	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

1.12 在经济学中, 所谓的 Lorenz 曲线是收入分布的图形表示. 请给出 Lorenz 曲线的一些描述, 并讨论 Lorenz 曲线与优势理论的联系.

第 2 章　均匀性度量

受总均值模型和著名的 Koksma-Hlawka 不等式的启发, 均匀试验设计的主要想法是把试验点均匀地散布在试验区域中. 这里, 均匀性测度是均匀试验设计的核心概念. 定义偏差为样本的经验分布函数和总体均匀分布函数之间的差异, 因此, 偏差是一种构造均匀设计的均匀性测度. 然而, 从不同角度出发, 可以定义不同的偏差. 本章介绍不同偏差的定义及其相应的下界, 其可用于构造均匀设计.

本章中, 2.1 节介绍总均值模型和 Koksma-Hlawka 不等式. 2.1—2.5 节介绍不同偏差的定义和性质, 其包括星偏差、中心化偏差、可卷偏差、混合偏差、离散偏差和 Lee 偏差等. 这些偏差大部分是用再生核希尔伯特空间这一工具来定义的. 2.6 节给出不同偏差的下界, 其可作为搜索均匀设计时的基准.

2.1　总均值模型

假设在区域 $\mathcal{X}$ 上某系统的真实模型为

$$y = f(\boldsymbol{x}), \tag{2.1}$$

其中 $\boldsymbol{x} = (x_1, \cdots, x_s)$ 为自变量或因子, y 为响应. 通常假设试验区域为一个超立方体 $[a_1, b_1] \times \cdots \times [a_s, b_s]$. 不失一般性, 假设其为单位超立方体 $C^s = [0,1]^s = [0,1] \times \cdots \times [0,1]$. 设 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 为 C^s 上的 n 个设计点. 一个自然的问题是基于 $\mathcal{P}$, 用近似模型

$$\hat{y} = g(\boldsymbol{x}) \tag{2.2}$$

来估计真实模型 $f(\boldsymbol{x})$. 通常, 有许多近似模型可用, 如何选择一个准则来评价这些近似模型是一个值得研究的问题. 我们可以先考虑最简单的总均值模型, 其目的是寻找 y 的总均值

$$E(y) = \int_{C^s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{2.3}$$

在一定意义下的一个最佳估计. 一个自然的想法是用 $\mathcal{P}$ 的样本均值

$$\bar{y}(\mathcal{P}) = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{2.4}$$

来估计 $E(y)$, 其中 $y_i = f(\boldsymbol{x}_i), i = 1, \cdots, n$. 给定设计点数 n, 在试验区域 $\mathcal{X}$ 寻找一个 n 个点的设计 $\mathcal{P}$，使之最小化两者之间的差异

$$\text{diff-mean} = |\bar{y}(\mathcal{P}) - E(y)|. \tag{2.5}$$

有两种方法选择 $\mathcal{P}$: *随机方法*和*确定性方法*. 随机方法的主要思想是寻找一个设计 $\mathcal{P}$ 使得样本均值 $\bar{y}(\mathcal{P})$ 是 $E(y)$ 的无偏估计或渐近无偏估计, 并具有最小的估计方差. 另一方面, 确定性方法是希望样本均值和总均值之间的差异 (2.5) 尽可能小.

最简单的随机方法是蒙特卡罗方法, 即 $\mathcal{P}$ 是 C^s 上的均匀分布 $U(C^s)$ 的 n 个独立样本, 其相应的样本均值 $\bar{y}(\mathcal{P})$ 是 $E(y)$ 的无偏估计, 且其估计方差为 $\text{Var}(f(\boldsymbol{x}))/n$, 其中随机变量 $\boldsymbol{x}$ 服从 C^s 上的均匀分布. 由中心极限定理可知, 样本均值和总均值之间的差异 95% 的置信区间为

$$|\bar{y}(\mathcal{P}) - E(y)| \leqslant 1.96\sqrt{\text{Var}(f(\boldsymbol{x}))/n}.$$

然而, 在很多情形下, 估计方差 $\text{Var}(f(\boldsymbol{x}))/n$ 太大了, 故需要减少. McKay 等 (1979) 提出的*拉丁超立方体抽样* (Latin hypercube sampling, LHS) 是一种被广泛使用的减少估计方差的方法. 该方法的主要思想为选取相关性的 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$, 但是它们具有相同的边缘分布. LHS 的构造方法是把试验区域 C^s 依每个 x_k 分成 n 层, 并具有相同的边缘概率 $1/n$, 然后对每个层抽样一次. 更细节的定义如下所示.

定义 2.1 *设试验区域为 $\mathcal{X} = [0,1]^s, \mathcal{U}$ 为全体 U-型设计 $U(n; n^s)$, 其每一列都是 $1/(2n), 3/(2n), \cdots, (2n-1)/(2n)$ 的一个置换. 拉丁超立方体设计 (Latin hypercube design, LHD) 或中点拉丁超立方体抽样 (midpoint Latin hypercube sampling, MLHS) 即 $\mathcal{U}$ 上的随机抽样. 等价地说, LHD 为一个 U-型设计 $U(n; n^s)$, 且其每一列都是 $1/(2n), 3/(2n), \cdots, (2n-1)/(2n)$ 的随机置换. 设 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 为拉丁超立方体设计, 每个 $\boldsymbol{x}_j$ 为一个边长为 $1/n$, 中心位于 $\boldsymbol{x}_j$ 的子立方体, 记为 $C_{\boldsymbol{x}_j}$. 设 $\boldsymbol{y}_j$ 为 $C_{\boldsymbol{x}_j}$ 上的随机样本, 则称集合 $\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n\}$ 为拉丁超立方体抽样.*

当 $f(\boldsymbol{x})$ 对每个变量 x_k 是单调的, McKay 等 (1979) 证明任何两个点之间具有负相关性. 记 LHS 的样本均值为 $\bar{y}_{\text{LHS}}$. Stein(1987) 和 Owen(1992a) 给出了 $\bar{y}_{\text{LHS}}$ 的方差的表达式, 且证明了

$$\text{Var}(\bar{y}_{\text{LHS}}) = \frac{\text{Var}(f(\boldsymbol{x}))}{n} - \frac{c}{n} + o\left(\frac{1}{n}\right),$$

其中 c 为一正常数. 因此, 该估计方差比蒙特卡罗方法得到的估计方差更低. 然而, LHS 并不能达到样本均值的方差的最小值. 一些改进方法可以比 LHS 的性能更好, 比如, Tang(1993) 提出的正交列拉丁超立方体设计; Koehler 和 Owen(1996) 给出一个综述性的讨论.

确定性方法是另一类常用的减少差异 (2.5) 的方法. 伪蒙特卡罗方法中著名的 Koksma-Hlawka 不等式说明

$$|\bar{y}(\mathcal{P}) - E(y)| \leqslant V(f)D^*(\mathcal{P}), \tag{2.6}$$

其中 $V(f)$ 为函数 f 在 Hardy 和 Krause 意义下的全变差, 具体定义见 (Niederreiter, 1992) 与 (Hua and Wang, 1981), 且 $D^*(\mathcal{P})$ 为 $\mathcal{P}$ 的不依赖于 f 的星偏差. 2.2 节将给出星偏差的定义和相应讨论. 在一些情形下, (2.6) 式中 Koksma-Hlawka 不等式的上界是紧上界. 给定函数 f 和试验区域, $V(f)$ 保持不变. 若 $V(f)$ 在试验区域上有界, 则可以选取试验区域 C^s 上的 n 个点的设计 $\mathcal{P}$, 使得其星偏差 $D^*(\mathcal{P})$ 尽可能小, 从而可以最小化 (2.5) 式的上界.

Weyl(1916) 提出的星偏差是伪蒙特卡罗方法中衡量均匀性的常用度量. 星偏差值越小, 点集的均匀性就越好. 因此, 从 Koksma-Hlawka 不等式可知, 可以选择一个有 n 个点的点集使其尽可能均匀地散布在试验区域 C^s 上. 称该点集为均匀设计 (uniform design, UD), 其由 Fang(1980) 与 Fang 和 Wang(1981) 提出. 均匀设计对模型是稳健的. 例如, 当两个函数 $f_1(\boldsymbol{x})$ 和 $f_2(\boldsymbol{x})$ 具有相同的全变差, 则均匀设计使得这两个函数在 (2.5) 的差异具有共同的上界.

需要指出的是, 拉丁超立方体抽样和均匀设计都是从总均值模型出发的. 在总均值模型意义下最好的设计可能不一定很好地估计在整个试验区域 C^s 上的真实模型 $f(\boldsymbol{x})$. 值得庆幸的是, 这两类设计都是空间填充设计, 其不仅可以用于估计总均值, 也可以得到 $f(\boldsymbol{x})$ 较好的估计.

2.2 星 偏 差

Weyl (1916) 提出的星偏差在均匀设计和伪蒙特卡罗方法中都有广泛应用. 星偏差可用于衡量 C^s 上的均匀概率分布和 C^s 上 n 个点相应的经验分布函数之间的差异.

2.2.1 定义

对任意 $\boldsymbol{x} = (x_1, \cdots, x_s) \in C^s$, 设 $F(\boldsymbol{x})$ 为 C^s 上均匀分布的分布函数, 即

$$F(\boldsymbol{x}) = \begin{cases} x_1 \cdots x_s, & 0 \leqslant x_i \leqslant 1,\ i = 1, \cdots, s, \\ 0, & \text{否则}. \end{cases} \tag{2.7}$$

设 $F_{\mathcal{P}}(\boldsymbol{x})$ 为设计 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}, \boldsymbol{x}_i \in C^s$ 的经验分布函数, 即

$$F_{\mathcal{P}}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} 1_{[\boldsymbol{x}_i, \boldsymbol{\infty})}(\boldsymbol{x}), \tag{2.8}$$

其中 $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{is}), \boldsymbol{\infty} = (\infty, \cdots, \infty)$, $1_{[\boldsymbol{x}_i, \boldsymbol{\infty})}(\boldsymbol{x})$ 为示性函数, 当 $\boldsymbol{x} \in [\boldsymbol{x}_i, \boldsymbol{\infty})$ 时, 其等于 1, 否则为 0, 且 $[\boldsymbol{x}_i, \boldsymbol{\infty}) = [x_{i1}, \infty) \times \cdots \times [x_{is}, \infty)$.

定义 2.2　*定义一个设计 $\mathcal{P}$ 的 L_p-星偏差为其 $F_{\mathcal{P}}(\boldsymbol{x})$ 和 $F(\boldsymbol{x})$ 之间差异的 L_p-范数, 即*

$$D_p^*(\mathcal{P}) = \| F_{\mathcal{P}}(\boldsymbol{x}) - F(\boldsymbol{x}) \|_p = \left\{ \int_{C^s} |F_{\mathcal{P}}(\boldsymbol{x}) - F(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \right\}^{1/p}. \tag{2.9}$$

从几何的观点来看, 经验分布函数可表示为 $F_{\mathcal{P}}(\boldsymbol{x}) = N(\mathcal{P} \cap [\boldsymbol{0}, \boldsymbol{x}))/n$, 其中 $N(\mathcal{P} \cap [\boldsymbol{0}, \boldsymbol{x}))$ 为设计 $\mathcal{P}$ 中落入超矩形 $[\boldsymbol{0}, \boldsymbol{x}) = [0, x_1) \times \cdots \times [0, x_s)$ 的设计点数. 均匀概率分布 $F(\boldsymbol{x})$ 为超矩形 $[\boldsymbol{0}, \boldsymbol{x})$ 的体积, 记为 $\mathrm{vol}([\boldsymbol{0}, \boldsymbol{x}))$. 则 (2.9) 式中的 L_p-星偏差可表示为

$$D_p^*(\mathcal{P}) = \left\{ \int_{C^s} \left| \frac{N(\mathcal{P} \cap [\boldsymbol{0}, \boldsymbol{x}))}{n} - \mathrm{vol}([\boldsymbol{0}, \boldsymbol{x})) \right|^p \mathrm{d}\boldsymbol{x} \right\}^{1/p}. \tag{2.10}$$

最常用的 p 值取为 $p = \infty$ 和 $p = 2$.

(1) 当 $p = \infty$ 时, L_∞-星偏差也简称为*星偏差*, 其可表示为

$$D^*(\mathcal{P}) = \sup_{\boldsymbol{x} \in C^s} \left| \frac{N(\mathcal{P} \cap [\boldsymbol{0}, \boldsymbol{x}))}{n} - \mathrm{vol}([\boldsymbol{0}, \boldsymbol{x})) \right|, \tag{2.11}$$

其恰好是拟合优度检验中的 Kolmogorov-Smirnov *统计量*. 星偏差不仅在伪蒙特卡罗方法中也在统计学中有重要应用, 不过星偏差并不能在多项式时间内计算其数值 (Winker and Fang, 1997).

(2) 当 $p = 2$ 时, Warnock (1972) 证明 L_2-星偏差有一个简单的表达式

$$\begin{aligned} [D_2^*(\mathcal{P})]^2 = & \left(\frac{1}{3}\right)^2 - \frac{2}{n} \sum_{i=1}^{n} \prod_{j=1}^{s} \frac{(1 - x_{ij}^2)}{2} \\ & + \frac{1}{n^2} \sum_{i,l=1}^{n} \prod_{j=1}^{s} [1 - \max(x_{ij}, x_{lj})], \end{aligned} \tag{2.12}$$

其计算复杂度为 $O(n^2 s)$. L_2-星偏差也是著名的 Cramér-Von Mises *拟合优度统计量*. 然而, 当 $p \neq 2$ 时, L_p-星偏差并没有显式表达式.

定义*局部偏差函数*为

$$\mathrm{disc}^*(\boldsymbol{x}) = \frac{N(\mathcal{P} \cap [\boldsymbol{0}, \boldsymbol{x}))}{n} - \mathrm{vol}([\boldsymbol{0}, \boldsymbol{x})), \tag{2.13}$$

其为点 $\boldsymbol{x}$ 处的均匀分布函数和经验分布函数之间的差异. 则 L_p-星偏差为

$$D_p^*(\mathcal{P}) = \|\mathrm{disc}^*\|_p = \begin{cases} \left\{ \displaystyle\int_{C^s} |\mathrm{disc}^*(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \right\}^{1/p}, & 1 \leqslant p < \infty, \\ \displaystyle\sup_{\boldsymbol{x} \in C^s} |\mathrm{disc}^*(\boldsymbol{x})|, & p = \infty. \end{cases} \tag{2.14}$$

当 $p < \infty$ 时, L_p-星偏差可以看成 C^s 上在 L_p-范数意义下局部偏差函数的平均值, 当 $p = \infty$ 时, L_p-星偏差可看成局部偏差函数绝对值的上确界. 图 2.1 显示在 C^s 上的两个设计和相应的局部偏差函数. 例如, 当 $\boldsymbol{x} = (0.5, 0.36)$, 设计 1 中 $\mathrm{disc}^*(\boldsymbol{x}) = 1/7 - 0.5 \times 0.36 = -0.0371$, 而对于设计 2, $\mathrm{disc}^*(\boldsymbol{x}) = 3/7 - 0.5 \times 0.36 = 0.2486$. 此外, 从局部偏差函数的波动可知, 第一个设计比第二个设计更均匀.

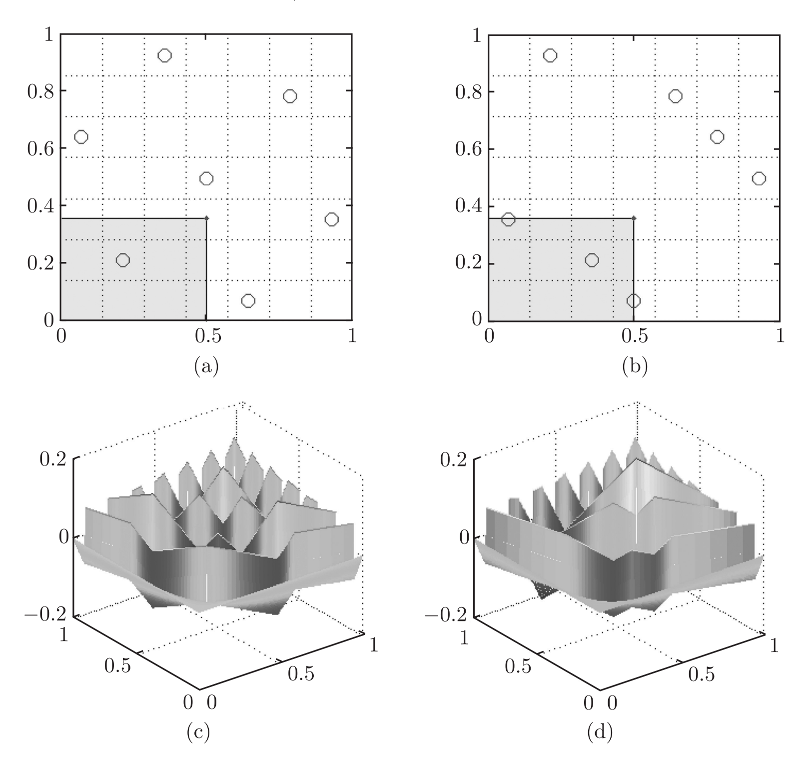


图 2.1 两个设计和其相应的局部偏差函数. (a) 设计 1, (b) 设计 2, (c) 设计 1 的局部偏差函数, (d) 设计 2 的局部偏差函数

2.2.2 性质

对于 C^s 上某一 n 个点的设计 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$, 定义 $\boldsymbol{X} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)^{\mathrm{T}}$ 为 $n \times s$ 设计矩阵, 其中 $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{is})^{\mathrm{T}}, i = 1, \cdots, n$. 对任一偏差 D, 定义 $D(\boldsymbol{X}) = D(\mathcal{P})$. 一个合理的偏差需要满足如下几个条件.

C_1 **对设计矩阵作行置换或列置换保持不变**. 显然, 对设计作行置换或列置换并不影响设计的性质, 因此, 均匀性测度也应该保持不变. 设 $\boldsymbol{P}: n \times n$ 和 $\boldsymbol{Q}: s \times s$ 为两个置换矩阵, 即其元素都是 0 或 1, 且每行和每列都只有一个 1. 若 D 满足

$$D(\boldsymbol{PXQ}) = D(\boldsymbol{X}), \quad \text{对任意设计矩阵}\boldsymbol{X},$$

称偏差 D 具有*置换不变性*. 这里, 对设计矩阵 $\boldsymbol{X}$ 左乘 $\boldsymbol{P}$ 相当于置换行向量, 而右乘 $\boldsymbol{Q}$ 相当于置换列向量.

C_2 **坐标旋转和反射保持不变**. 直观上说, 当一个设计关于某一坐标旋转或反射时, 应该不影响设计的均匀性. 设 $\boldsymbol{Q}_\theta$ 为 $x_i - x_j$ 平面上的 $s \times s$ 旋转矩阵. 例如, 在 $x_1 - x_2$ 平面上顺时针旋转 θ 角度的旋转矩阵为

$$\boldsymbol{Q}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & & & \\ \sin\theta & \cos\theta & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}. \tag{2.15}$$

因此, 在 $x_i - x_j$ 平面上旋转设计 $\boldsymbol{X}$ 可表示为

$$\boldsymbol{X}_\theta = (\boldsymbol{X} - 0.5)\boldsymbol{Q}_\theta + 0.5, \tag{2.16}$$

其中对于某一矩阵 $\boldsymbol{A} = (a_{ij})$, 定义 $\boldsymbol{A} \pm 0.5 = (a_{ij} \pm 0.5)$. 称偏差 D 为*旋转不变偏差*, 若其满足

$$D(\boldsymbol{X}_\theta) = D(\boldsymbol{X}), \quad 当\ \theta = \pi/2, \pi, 3\pi/2, 且对于任意设计矩阵\boldsymbol{X}.$$

进一步地, 定义设计 $\mathcal{P}$ 关于平面 $x_j = 1/2$ 的反射为

$$\mathcal{P}_{\mathrm{ref},j} = \{(x_{i1}, \cdots, x_{i,j-1}, 1 - x_{ij}, x_{i,j+1}, \cdots, x_{is})^{\mathrm{T}} : i = 1, \cdots, n\}.$$

称偏差 D 为*反射不变偏差*, 若对任意设计 $\mathcal{P}$,

$$D(\mathcal{P}_{\mathrm{ref},j}) = D(\mathcal{P}), \quad j = 1, \cdots, s.$$

易知, 在 $x_i - x_j$ 平面上顺时针旋转 $\pi/2, \pi$ 或 $3\pi/2$ 的角度, 且 $[0,1]^s$ 对平面 $x_j = 1/2$ 进行反射, 试验区域保持不变. 旋转不变偏差和反射不变偏差要求这些旋转和反射变换不会影响设计的均匀性.

C_3 **可以衡量投影均匀性**. 根据 1.3.2 小节的等级排序原则, 低阶效应比高阶效应更重要, 同阶效应同等重要. 因此, 需要考虑设计在每一维上的投影均匀性以估计主效应, 以及 $x_i - x_j$ 平面上的投影均匀性以估计因子 i 和因子 i 之间的交互效应, 等等. 一个好的均匀性度量, 不仅能衡量设计 $\mathcal{P}$ 在整个试验区域 C^s 上的均匀性, 也能衡量 $\mathcal{P}$ 在 C^u 上的投影均匀性, 其中 u 是 $\{1, \cdots, s\}$ 的非空集合.

C_4 **有几何意义**. 均匀设计有其几何意义, 因此均匀性度量也需要有几何意义.

C_5 **容易计算**. 给定某一均匀性度量, 通常并不容易搜索一个在该均匀性度量下的均匀设计. 一般地, 我们采用一些优化算法用于寻找均匀设计. 在搜索过程中, 需要计算每个设计的度量值, 因此, 我们希望均匀性度量值可以在关于设计点数 n 的多项式时间内得到.

C_6 **满足 Koksma-Hlawka 不等式**. 不等式 (2.6) 提供了均匀设计的一个重要理论支撑, 因此每一个均匀性度量都应该满足该不等式.

现考虑 (2.11) 式中的星偏差和 (2.12) 式中的 L_2-星偏差的性质. 易知, 这两个偏差都符合 C_1, C_4 和 C_6. 对于其他准则, 有下面的讨论.

对于 C_2. 星偏差和 L_2-星偏差都不满足 C_2. 当对设计 $\mathcal{P}$ 进行旋转和反射变换后, 这两个偏差值都可能会改变. 其改变的原因在于 L_p-星偏差的定义中包含 $[\mathbf{0},\boldsymbol{x})$, 因此原点 $\mathbf{0}$ 起到关键作用. 我们通过一个例子来说明星偏差的这种性质. 考虑对于正方形 $\boldsymbol{C}^2$ 上的设计进行如 (2.15) 式的顺时针旋转

$$\boldsymbol{Q}_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\quad \theta=\pi/2,\pi,3\pi/2,$$

并对反射轴 $x_1=1/2$ 和 $x_2=1/2$ 作反射变换. 图 2.2 显示旋转和反射变换后星偏差值会变化, 因此星偏差既不具有旋转不变性也不具有反射不变性. 需要指出的是, 图 2.2(b) 中旋转后的设计和图 2.2(e) 中反射变换后的设计有相同的星偏差, 且图 2.2(d) 的设计和图 2.2(f) 的设计也有同样的星偏差, 因为这些设计可以通过关于直线 $y=x$ 的反射而相互得出. L_2-星偏差有类似的性质.

对于 C_3. 除 $p=\infty$ 之外, L_p-星偏差不能衡量低维投影设计的均匀性, 即星偏差满足 C_3, 而 L_2-星偏差不能满足. 从定义 (2.14) 可知, 当 $p<\infty$ 时, L_p-星偏差的局部偏差函数 $\mathrm{disc}^*(\boldsymbol{x})$ 忽略低维子区域.

设 $u\subseteq\{1,\cdots,s\}$ 为感兴趣因子的指标集, $[0,1]^u$ 为相应的试验区域. 定义限制在指标集 u 中的因子的局部偏差函数为

$$\mathrm{disc}_u^*(\boldsymbol{x})\triangleq\mathrm{disc}^*\left(\boldsymbol{x}_u,\mathbf{1}\right),\tag{2.17}$$

其中 $\boldsymbol{x}_u$ 为 $\boldsymbol{x}$ 在 $[0,1]^u$ 上的投影, 且 $\mathrm{disc}^*(\boldsymbol{x})$ 如 (2.13) 定义. 例如, 对于 $\boldsymbol{x}=(x_1,\cdots,x_s)$, $(\boldsymbol{x}_u,\mathbf{1})=(1,x_2,1)$, 当 $s=3$ 且 $u=\{2\}$; 而 $(\boldsymbol{x}_u,\mathbf{1})=(x_1,1,x_3,1)$, 当 $s=4$ 且 $u=\{1,3\}$. 需要指出的是 $\mathrm{disc}_\varnothing^*(\boldsymbol{x})=0$. 限制在指标集 u 的投影设计的 L_p-星偏差可记为 $D_{p,u}^*(\mathcal{P})=\|\mathrm{disc}_u^*\|_p$. 星偏差满足

$$D^*(\mathcal{P})=\sup_u D_{\infty,u}^*(\mathcal{P}).$$

因此, 星偏差已考虑了投影均匀性.

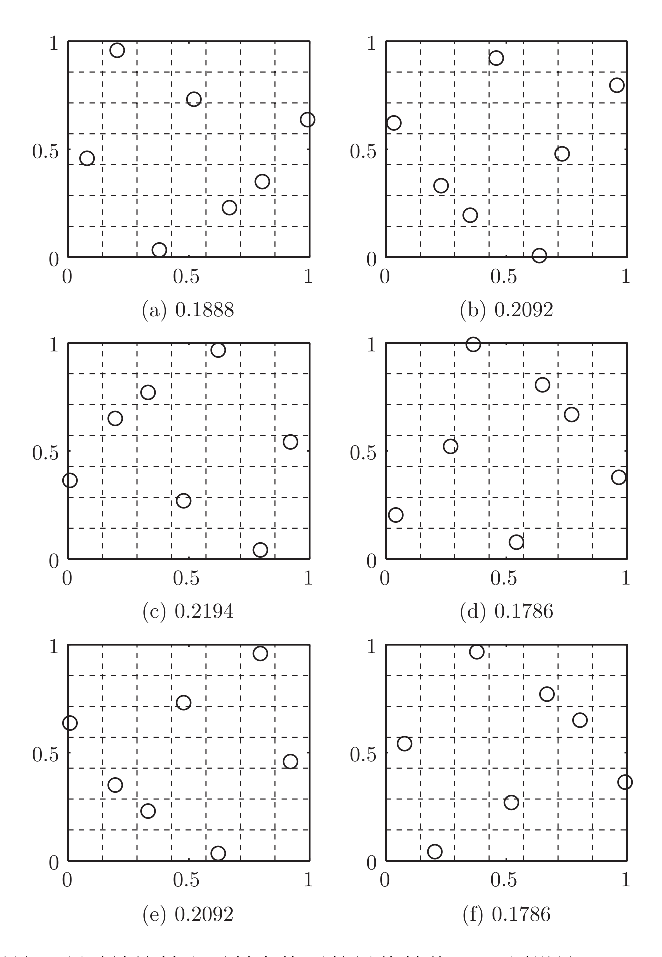


图 2.2 某设计及顺时针旋转和反射变换后的星偏差值. (a) 原设计, (b)—(d) 分别旋转 $\pi/2, \pi, 3\pi/2$ 角度后的设计, (e) 对 $x_1 = 1/2$ 作反射变换, (f) 对 $x_2 = 1/2$ 作反射变换

对于 C_5. L_2-星偏差有显式表达式, 而星偏差没有, 因此只有 L_2-星偏差满足 C_5. 为了计算星偏差值, 需要计算当 $\boldsymbol{x}$ 以任何方向趋于集合 $\{0,1,x_{11},\cdots,x_{n1}\}\times\cdots\times\{0,1,x_{1s},\cdots,x_{ns}\}$ 时, $|\mathrm{disc}^*(\boldsymbol{x})|$ 的极限值, 因为局部偏差函数是 $\boldsymbol{x}$ 的逐段多元线性函数, 如图 2.1 所示. Winker 和 Fang (1997) 指出, 这需要 $\mathrm{O}(n^s)$ 运算. 因此, 即使对于中等大小的试验次数, 也难以计算星偏差值. 为此, 人们考虑用快速有效的方法计算不同维数的星偏差的近似值, 例如, Niederreiter(1973) 考虑一维情形, Clerk (1986) 考虑二维情形, Bundschuh 和 Zhu(1993) 考虑 $s \leqslant 10, n \leqslant 100$, Winker 和 Fang (1997) 考虑 n 比较大的情形.

总之, 星偏差满足 C_1, C_3, C_4 和 C_6, 而 L_2-星偏差满足 C_1, C_4, C_5 和 C_6. 此外, 当 $p \neq 2$ 且 $p < +\infty$ 时, L_p-星偏差没有显式表达式且不能衡量投影均匀性, 即

其仅满足 C_1, C_4 和 C_6. 2.3 节将提出一些广义 L_2-偏差, 其可以克服 L_p-星偏差的缺点.

2.3 广义 L_2-偏差

为了克服 L_p-星偏差的缺点, 本节将介绍 Hickernell (1998a, 1998b) 和 Zhou 等 (2013) 提出的一些广义 L_2-偏差, 简称为广义偏差.

由于 L_2-星偏差有显式表达式, 广义的偏差也考虑 L_2-范数. 由于 L_2-星偏差不能度量投影均匀性, 广义偏差应该可以度量任意低维投影区域上的均匀性. 此外, 原点 $\mathbf{0}$ 在计算 L_2-星偏差时具有关键作用, 因此导致一些不合理的现象. 广义偏差也需要消除原点 $\mathbf{0}$ 的特殊影响, 即使得超矩形的每个顶点与原点 $\mathbf{0}$ 具有同等重要.

2.3.1 定义

设 $\mathcal{P}$ 为 C^s 上一个具有 n 个设计点, s 个因子的设计. 记 $\{1:s\}=\{1,\cdots,s\}$, $u\subset\{1:s\}$ 为感兴趣因子的指标集, $[0,1]^u$ 为由指标集 u 对应的超矩形, $\boldsymbol{x}_u=(x_j)_{j\in u}$ 为 $\boldsymbol{x}$ 在 $[0,1]^u$ 上的投影, $\mathcal{P}_u$ 为设计 $\mathcal{P}$ 在 $[0,1]^u$ 上的投影. 对任意 $\boldsymbol{x}_u\in[0,1]^u$, 记 $R_u(\boldsymbol{x}_u)\subseteq[0,1]^u$ 为预先确定的区域. 本小节的最后将给出不同偏差对应的 $R_u(\boldsymbol{x}_u)$. 由指标集 u 确定的因子的*局部投影偏差*定义如下:

$$\operatorname{disc}_u^R(\boldsymbol{x}_u)=\operatorname{Vol}(R_u(\boldsymbol{x}_u))-\frac{|\mathcal{P}_u\cap R_u(\boldsymbol{x}_u)|}{n},\tag{2.18}$$

其可认为是 $\boldsymbol{x}_u$ 的函数. 定义 disc_u^R 的 L_2-范数如下:

$$\left\|\operatorname{disc}_u^R\right\|_2=\left\{\int_{[0,1]^u}\left|\operatorname{disc}_u^R(\boldsymbol{x}_u)\right|^2\mathrm{d}\boldsymbol{x}_u\right\}^{1/2}=\left\{\int_{[0,1]^u}\left|\operatorname{disc}_u^R(\boldsymbol{x}_u)\right|^2\mathrm{d}\boldsymbol{x}\right\}^{1/2}.$$

定义广义 L_2-偏差如下:

$$\begin{aligned}D_2^R(\mathcal{P})=&\left\|\left(\left\|\operatorname{disc}_u^R\right\|_2\right)_{u\subseteq\{1:s\}}\right\|_2\\=&\left\{\sum_{u\subseteq\{1:s\}}\int_{[0,1]^u}\left|\operatorname{disc}_u^R(\boldsymbol{x}_u)\right|^2\mathrm{d}\boldsymbol{x}_u\right\}^{1/2}\\=&\left\{\int_{[0,1]^s}\sum_{u\subseteq\{1:s\}}\left|\operatorname{disc}_u^R(\boldsymbol{x}_u)\right|^2\mathrm{d}\boldsymbol{x}\right\}^{1/2}.\end{aligned}\tag{2.19}$$

其考虑所有的局部投影偏差. 显然, (2.19) 式中的广义 L_2-偏差考虑投影均匀性, 因此满足 C_3. 当因子个数从 s 增长到 $\tilde{s}$, 偏差可能会增大, 因为其增加了 $\left\|\mathrm{disc}_u^R\right\|_2$ 这一部分, 其中 u 包括 $\{(s+1):\tilde{s}\}$ 的所有非空子集.

进一步地, 为了满足 C_4, $R_u(\boldsymbol{x}_u)$ 的定义需要有直观的几何解释. 在 (2.18) 式中不同的 $R_u(\boldsymbol{x}_u)$ 导致不同的广义偏差. Hickernell (1998a, 1998b) 提出多个广义 L_2-偏差, 其中*中心化偏差* (centered L_2-discrepancy, CD) 和*可卷偏差* (wrap-around L_2-discrepancy, WD) 有很多理论研究和实际应用. 这两个偏差都可以克服 L_p-星偏差的一些缺点. 不过, Zhou 等 (2013) 指出这两个偏差也存在缺点, 并提出一种新的偏差, 称为*混合偏差* (mixture discrepancy, MD). 在下面的小节中, 将详细给出不同 $R_u(\boldsymbol{x}_u)$ 的定义以及中心化偏差、可卷偏差和混合偏差的表达式.

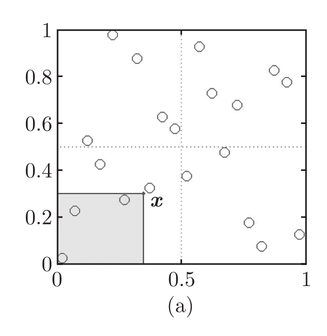
2.3.2　中心化偏差

记单位立方体 $[0,1]^s$ 的 2^s 个顶点为 $\{0,1\}^s$. 对于任意点 $\boldsymbol{x}\in[0,1]^s$, 记 $\boldsymbol{a}_{\boldsymbol{x}}\in\{0,1\}^s$ 为最靠近点 $\boldsymbol{x}$ 的顶点, 即 $\boldsymbol{a}_{\boldsymbol{x}}=(a_{x_1},\cdots,a_{x_s})$, 其中当 $0\leqslant x_j\leqslant 1/2$ 时, $a_{x_j}=0$, 而当 $1/2<x_j\leqslant 1$ 时, $a_{x_j}=1$. 对任意投影 $u\subset\{1:s\}$, 记 $a_{\boldsymbol{x}_u}$ 为 $[0,1]^u$ 中最靠近 $\boldsymbol{x}_u$ 的顶点. 定义 $R_u^C(\boldsymbol{x}_u)$ 为 $\boldsymbol{a}_{\boldsymbol{x}_u}$ 和 $\boldsymbol{x}_u$ 之间的超矩形.

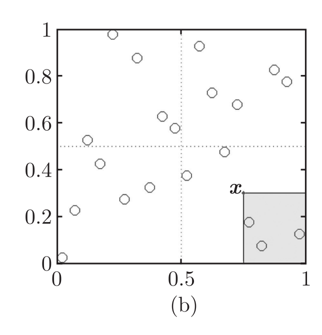
对任意 $\boldsymbol{x},\boldsymbol{y}\in[0,1]^s$, 设 $J(\boldsymbol{x},\boldsymbol{y})$ 为包含 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的最小超矩形, 即

$$J(\boldsymbol{x},\boldsymbol{y})=\{(t_1,\cdots,t_s):\min(x_j,y_j)\leqslant t_j\leqslant\max(x_j,y_j),\ j=1,\cdots,s\}.$$

若 $\boldsymbol{x}$ 的元素都不大于 $\boldsymbol{y}$ 的相应位置元素, 则 $J(\boldsymbol{x},\boldsymbol{y})=[\boldsymbol{x},\boldsymbol{y}]$. 对于中心化偏差, 定义 $R_u^C(\boldsymbol{x}_u)=J(\boldsymbol{a}_{\boldsymbol{x}_u},\boldsymbol{x}_u)$. 例如, 当 $u=\{1,2\}$ 时, 单位正方形 $[0,1]^2$ 由中心点而分为四个小的正方形, 每个小正方形包括一个 $[0,1]^2$ 的顶点和中心点 $(1/2,1/2)$. 因此, $\boldsymbol{a}_{\boldsymbol{x}_u}$ 和 $\boldsymbol{x}_u$ 为超矩形 $R_u^C(\boldsymbol{x}_u)$ 的两个对角顶点. 图 2.3 给出中心化偏差对于 $u=\{1,2\}$ 的局部投影偏差.



(a)



(b)

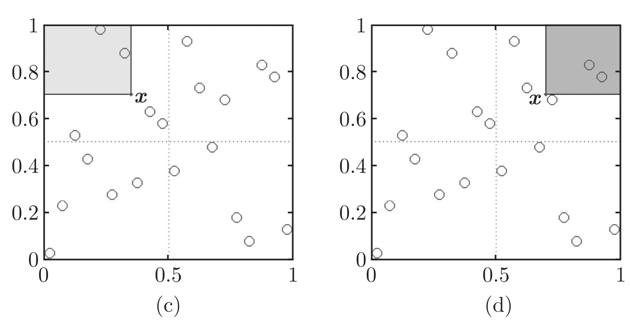


图 2.3 中心化偏差中的 $R^C_{\{1,2\}}(\boldsymbol{x})$

对于中心化偏差, (2.18) 式的局部投影偏差变为

$$\begin{aligned}\operatorname{disc}^C_u(\boldsymbol{x}_u)=&\operatorname{Vol}(R^C_u(\boldsymbol{x}_u))-\frac{\left|\mathcal{P}_u\cap R^C_u(\boldsymbol{x}_u)\right|}{n}\\=&\operatorname{Vol}(J(\boldsymbol{a}_{\boldsymbol{x}_u},\boldsymbol{x}_u))-\frac{|\mathcal{P}_u\cap J(\boldsymbol{a}_{\boldsymbol{x}_u},\boldsymbol{x}_u)|}{n}.\end{aligned}\tag{2.20}$$

由 (2.19), 中心化偏差变为

$$\operatorname{CD}(\mathcal{P})=\left\|\left(\left\|\operatorname{disc}^C_u\right\|_2\right)_{u\subseteq\{1:s\}}\right\|_2.$$

Hickernell(1998b) 证明中心化偏差有下面的解析表达式:

$$\begin{aligned}\operatorname{CD}(\mathcal{P})=&\left\{\left(\frac{13}{12}\right)^s-\frac{2}{n}\sum_{i=1}^n\prod_{j=1}^s\left(1+\frac{1}{2}|x_{ij}-0.5|-\frac{1}{2}|x_{ij}-0.5|^2\right)\right.\\&\left.+\frac{1}{n^2}\sum_{i,k=1}^n\prod_{j=1}^s\left(1+\frac{1}{2}|x_{ij}-0.5|+\frac{1}{2}|x_{kj}-0.5|-\frac{1}{2}|x_{ij}-x_{kj}|\right)\right\}^{1/2}.\end{aligned}\tag{2.21}$$

显然, 中心化偏差满足准则 C_1, C_3, C_4 和 C_5. 由于区域 $R^C_u(\boldsymbol{x}_u)$ 把原点 $\mathbf{0}$ 和 2^s 个顶点等同来看, 因此对设计进行反射和旋转变换不改变中心化偏差值, 因此, 中心化偏差具有旋转不变性和反射不变性, 即满足准则 C_2. Hickernell(1998b) 证明了中心化偏差也满足 Koksma-Hlawka 不等式 (C_6), 不过其中全变差的定义需要修改.

2.3.3 可卷偏差

从定义可知, L_p-星偏差和中心化偏差都要求 $R_u(\boldsymbol{x}_u)$ 的定义中包含单位立方体的一个或更多的顶点. 一个对预先定义的区域 $R_u(\boldsymbol{x}_u)$ 自然的推广是使其落入单位超立方体内部, 且不包含单位超立方体的任何顶点, 因此, 可以定义其为 $[0,1]^u$

中两点 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 之间的超矩形区域. 此外, 为了满足准则 C_2, 使超立方体的每一维坐标都有可卷性质, 超立方体像一个圆环, 即把 $x_j = 0$ 和 $x_j = 1$ 视为同一个点. 区域 $R_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u)$ 和局部偏差函数分别定义如下:

$$R_j^W(y_j, x_j) = \begin{cases} [y_j, x_j], & y_j \leqslant x_j, \\ [0, x_j] \cup [y_j, 1], & x_j < y_j, \end{cases}$$

$$R_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u) = \bigotimes_{j \in u} R_j^W(y_j, x_j),$$

$$\mathrm{disc}_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u) = \mathrm{Vol}(R_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u)) - \frac{|\mathcal{P} \cap R_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u)|}{n}, \tag{2.22}$$

其中 $\bigotimes$ 为 Kronecker 乘积. 图 2.4 显示一些区域 $R_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u)$, 其中两个阴影的矩形视为连在一起的同一区域.

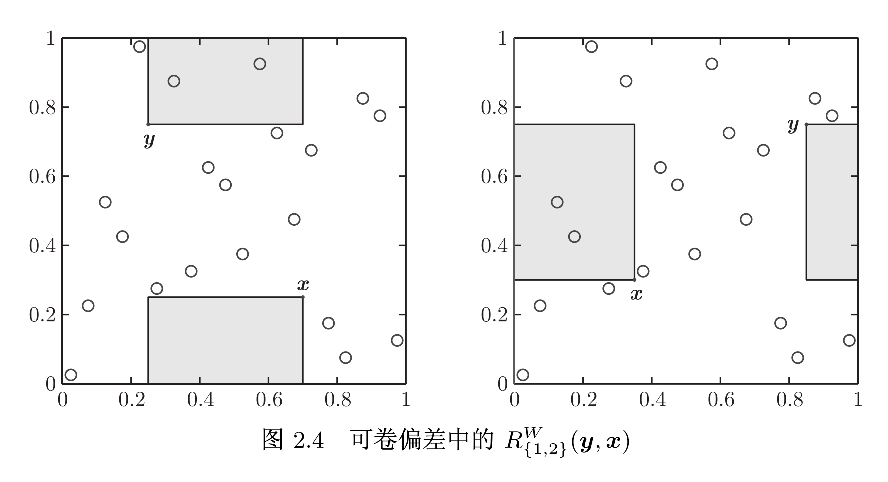


图 2.4 可卷偏差中的 $R_{\{1,2\}}^W(\boldsymbol{y}, \boldsymbol{x})$

由于区域具有可卷性质, 称相应的偏差为**可卷偏差**. 由 (2.19) 式, 定义可卷偏差为

$$\mathrm{WD}(\mathcal{P}) = \left(\sum_{u \subseteq \{1:s\}} \int_{[0,1]^{2u}} \left| \mathrm{disc}_u^W(\boldsymbol{y}_u, \boldsymbol{x}_u) \right|^2 \mathrm{d}\boldsymbol{x}_u \mathrm{d}\boldsymbol{y}_u \right)^{1/2}.$$

可卷偏差的平方值可表示为 (Hickernell, 1998a)

$$\mathrm{WD}^2(\mathcal{P}) = -\left(\frac{4}{3}\right)^s + \frac{1}{n^2} \sum_{i,k=1}^{n} \prod_{j=1}^{s} \left[\frac{3}{2} - |x_{ij} - x_{kj}| + |x_{ij} - x_{kj}|^2 \right]. \tag{2.23}$$

易知, 可卷偏差满足准则 C_1, C_3, C_4 和 C_5. 由于可卷偏差的可卷性质消除了原点 $\mathbf{0}$ 的影响, 则准则 C_2 也满足, 即可卷偏差也具有旋转不变性和反射不变性.

Hickernell (1998a) 证明可卷偏差也满足准则 C_6, 其中 Koksma-Hlawka 不等式中的全变差也需要一些修改.

2.3.4 可卷偏差和中心化偏差的一些讨论

我们已经说明了中心化偏差和可卷偏差都满足准则 C_1—C_6. 因此, 在文献中构造均匀设计时大量采用这两个偏差. 网站上已有的均匀设计表大部分也基于这两个偏差. 然而, 下面通过两个例子来说明在中心化偏差准则下, 在某些情形下会出现一些与直观不相符的现象.

例 2.1 考虑下面的两个 U-型设计 $U(24;3^3)$

$$\boldsymbol{X}_1=\begin{bmatrix} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3 \\ 1\ 1\ 2\ 2\ 2\ 3\ 3\ 3\ 1\ 1\ 1\ 2\ 2\ 2\ 3\ 3\ 1\ 1\ 1\ 2\ 2\ 3\ 3\ 3 \\ 1\ 3\ 1\ 2\ 3\ 1\ 2\ 3\ 1\ 2\ 3\ 1\ 2\ 3\ 1\ 2\ 1\ 2\ 3\ 2\ 3\ 1\ 2\ 3 \end{bmatrix}^{\mathrm{T}}$$

和

$$\boldsymbol{X}_2=\begin{bmatrix} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3 \\ 1\ 1\ 1\ 2\ 2\ 3\ 3\ 3\ 1\ 1\ 2\ 2\ 2\ 2\ 3\ 3\ 1\ 1\ 1\ 2\ 2\ 3\ 3\ 3 \\ 1\ 2\ 3\ 1\ 3\ 1\ 2\ 3\ 1\ 3\ 2\ 2\ 2\ 2\ 1\ 3\ 1\ 2\ 3\ 1\ 3\ 1\ 2\ 3 \end{bmatrix}^{\mathrm{T}},$$

其中 $\boldsymbol{A}^{\mathrm{T}}$ 表示矩阵 $\boldsymbol{A}$ 的转置. 考虑从三水平三因子的完全因子设计的 27 个试验次数中无放回地选 24 个试验点组成一个子设计, 则共有 $\begin{pmatrix} 27 \\ 24 \end{pmatrix}=2925$ 种不同的子设计. 可知, $\boldsymbol{X}_1$ 在这 2925 个设计中具有最低的中心化偏差和最低的可卷偏差, 而且该设计恰好为无重复点的 U-型设计. $\boldsymbol{X}_2$ 来自均匀设计网站

$$\text{http://www.math.hkbu.edu.hk/UniformDesign/}$$

在中心化偏差意义下的最佳设计. 简单计算后可知, 在中心化偏差意义下, $\boldsymbol{X}_2$ 比 $\boldsymbol{X}_1$ 更均匀, 因为

$$\mathrm{CD}^2(\boldsymbol{X}_1)=0.032779,\quad \mathrm{CD}^2(\boldsymbol{X}_2)=0.032586.$$

然而, $\mathrm{WD}^2(\boldsymbol{X}_1)=0.100852$, $\mathrm{WD}^2(\boldsymbol{X}_2)=0.101732$, 因此, 在可卷偏差意义下, $\boldsymbol{X}_1$ 比 $\boldsymbol{X}_2$ 更均匀. 这个例子说明在不同均匀性度量下, 均匀设计可能不一样. 此外, 在 $\boldsymbol{X}_2$ 的 24 个试验点里中心点 (2 2 2) 重复了 4 次. 我们知道在计算机试验中, 重复点并不能带来更多信息, 因此计算机试验要求设计没有重复点. 均匀设计是计算机试验的一种重要设计方法, 则设计 $\boldsymbol{X}_2$ 并不合理. Zhou 等 (2013) 指出, 出现这种情况的原因在于中心化偏差准则对位于中心点附近的区域关注比较少, 这可以从

其定义中的 $R_u^C(\boldsymbol{x}_u)$ 得知, 因此该偏差对中心点附近的设计点是否散布均匀关注较少. 这种缺点在高维情形中更加凸显, 则中心化偏差并不适合于构造高维情形下的均匀设计.

下面考虑中心化偏差的维数效应问题. 设 $\mathcal{P}_0 = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 为单位超立方体 C^s 上的 n 个设计点, 其由中点拉丁超立方体抽样而得到 (Fang et al., 2006a). 设

$$\mathcal{Y}(r) = \{\boldsymbol{y}_j = r(\boldsymbol{x}_j - \boldsymbol{m}) + \boldsymbol{m}, j = 1, \cdots, n\}, \tag{2.24}$$

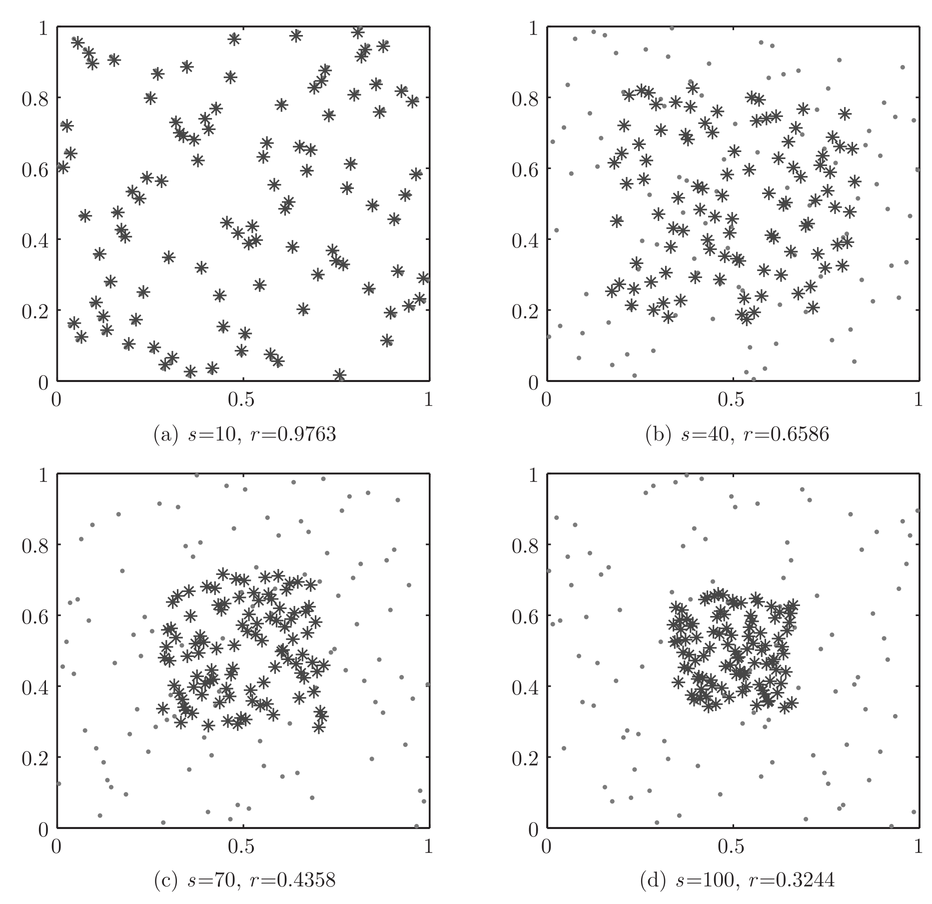
其中 $\boldsymbol{m} = \dfrac{1}{2}\mathbf{1}_s$ 为 C^s 的中心点且 $0 < r < 1$. 当比例 r 从 1 降为 0 时, $\boldsymbol{y}_j$ 沿着 $\boldsymbol{x}_j$ 和 $\boldsymbol{m}$ 之间的方向, 从点 $\boldsymbol{x}_j$ 收敛到中心点 $\boldsymbol{m}$. 在 $0 < r \leqslant 1$ 范围内最小化点集 $\mathcal{Y}(r)$ 的中心化偏差, 并记 $\mathcal{P}_1$ 为其中具有最小中心化偏差值的点集. 对于不同维数, 表 2.1 列出 100 次重复中点拉丁超立方体抽样而得到的 $\mathcal{P}_1$ 相应的 100 个 r 的平均值以及标准差.

表 2.1　100 次最佳设计 $\mathcal{P}_1$ 相应的 r 的平均值和标准差

s	10	20	30	40	50	60	70	80	90	100
$\mathcal{P}_1 : \bar{r}$	0.974	0.894	0.775	0.660	0.569	0.495	0.440	0.390	0.352	0.320
(std)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.005)	(0.002)	(0.000)	(0.004)	(0.000)

表 2.1 显示维数从 10 增大至 100 时, $\mathcal{P}_1$ 相应的比例 r 的平均值 $\bar{r}$ 会显著减小, 且 r 的标准差也会趋于 0. 为了更直观地说明, 考虑对于 $s = 10, 40, 70, 100$, 分别由中点拉丁超立方体抽样方法 (见 (Fang et al., 2006a)) 产生一个设计 $\mathcal{P}_0$. 图 2.5 给出 $\mathcal{P}_0$ 和相应的 $\mathcal{P}_1$ 的前两维投影结果. 由其可知, 随着因子个数的增大, 在中心化偏差下的最优比例 r 会减小. 除了图 2.5 中给出的前两维的投影图, 所有 $\begin{pmatrix} s \\ 2 \end{pmatrix}$ 个二维投影图都是类似的. 特别地, 当 $s = 100$ 时, 某一个拉丁超立方体设计 $\mathcal{P}_0$ 和相应 $\mathcal{P}_1$ 的中心化偏差分别为 56411746.8 和 2962.0, 这说明在该偏差意义下 $\mathcal{P}_1$ 比 $\mathcal{P}_0$ 更均匀. 在中心化偏差意义下, 随着因子个数的增大, 最优收缩比例减小, 集中在中心点的程度加剧. 因此, 作为均匀性度量, 中心化偏差具有严重的维数祸根问题. 在这种意义下, 中心化偏差并不是一个理想的均匀性度量, 尤其对于高维情形.

Fang 和 Ma (2001b) 指出, 对于一维设计而言, 位置平移不改变其可卷偏差值. 实际上, 从可卷偏差的表达式 (2.23) 可知, 任意如 (2.25) 式的水平平移都不改变其可卷偏差值. 然而, 这种灵活性可能会带来一些不合理的现象. 图 2.6 给出的四个设计具有相同的可卷偏差值, 不过直观上来看, 设计 D_1 更均匀一点. 导致这种现象的原因在于可卷偏差定义中的可卷性, 其把每一维的 0 和 1 看为同一个点. 因此, 可卷偏差对一个或多个因子的水平平移不敏感. 下面给出评价均匀性度量的更多准则.



(a) s=10, r=0.9763　(b) s=40, r=0.6586

(c) s=70, r=0.4358　(d) s=100, r=0.3244

图 2.5　在中心化偏差下, 不同维数的最优收缩比例, 其中
“.” 和 “$*$” 分别表示 $\mathcal{P}_0$ 和 $\mathcal{P}_1$ 的前两维投影点

C_7 **对一维或多维水平平移具有敏感性**. 一个均匀性度量 D 需要对小的水平平移具有敏感性, 即水平平移会改变其取值.

考虑设计 $\mathcal{P}$ 的一维或多维的水平平移:

$$\mathcal{P}' = \{(x_{i1} + a_1, \cdots, x_{is} + a_s)^{\mathrm{T}} \pmod 1),\ i = 1, \cdots, n\}, \tag{2.25}$$

其中 $a_j \in [0, 1),\ j = 1, 2, \cdots, s$, 算子 (mod 1) 表示每个元素都模 1 使得 $\mathcal{P}'$ 的设计点都落入 $[0, 1]^u$. 需要指出的是, 该准则并不要求偏差 D 对任意的 $a_j \in [0, 1) (j = 1, 2, \cdots, s)$ 都会改变.

C_8 **没有维数祸根问题**. 我们知道许多问题在低维和高维的表现是截然不同的, 即同一问题从低维变为高维时其性质发生了显著变化. 这种现象被称为维数祸根问题. 很多均匀性度量都与体积有强相关, 因此需要研究这些准则在高维的表现.

根据均匀分布的要求, 单位立方体中每一点的地位应该相同.

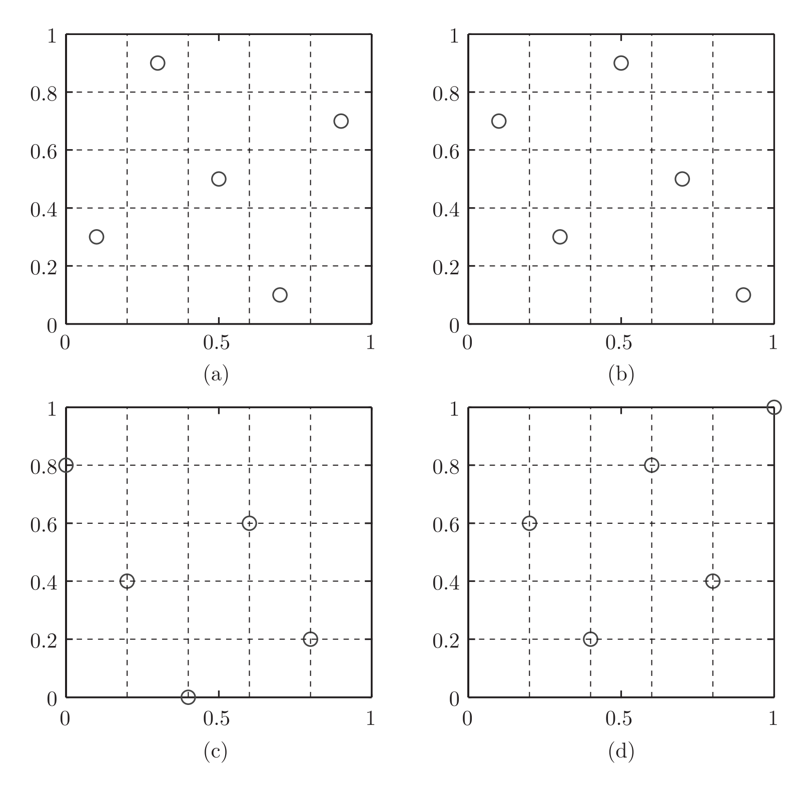


图 2.6　由水平平移得到的四个设计具有相同的可卷偏差值

下面检查 L_p-星偏差是否满足准则 C_7 和 C_8.

对于 C_7. 当设计点在一维或多维的某个方向上平移时, 如 (2.25) 所示, 给定 $\boldsymbol{x}$, 定义 (2.10) 或 (2.11) 中 $N(\mathcal{P}\cap[\mathbf{0},\boldsymbol{x}])$ 的取值会有所变化, 因此星偏差或 L_2-星偏差的取值也可能变化. 因此, 星偏差和 L_2-星偏差对一维或多维上的平移会敏感.

对于 C_8. 在高维情形下, L_p-星偏差会出现一些不合理现象. 例如, 在星偏差意义下, 考虑一个点的均匀设计, 则该点为 $\mathcal{P}=\{(z,\cdots,z)\}$, 其中 z 满足 $z^s+z-1=0$. 因此当 $s=1$ 时, $z=1/2$; 当 $s=2$ 时, $z=(\sqrt{5}-1)/2$, 且随着 s 的增大, z 趋于 1. 然而, 从直观上来说, 对于任意 s, 一个点的均匀设计应该位于 C^s 的中心, 即不管 s 取多大, $z=1/2$. 出现这种现象的原因在于 L_p-星偏差中原点 $\mathbf{0}$ 占据特殊位置有关.

对于中心化偏差, 从图 2.5 和表 2.1 可知, 该偏差并不满足 C_8, 但满足 C_7. 例如, 考虑例 2.1 的 $\boldsymbol{X}_1$, 设 $\boldsymbol{U}=(\boldsymbol{X}_1-0.5)/3$ 为其导出矩阵. 在 (2.25) 式中设 $(a_1,a_2,a_3)=(0.1,0,0)$, 即仅平移该设计的第一维. 则平移后的设计的中心化偏差

的平方值为 0.0443, 其与 $\boldsymbol{U}$ 的不同.

对于可卷偏差, 易知其不满足 C_7, 因为对于 (2.25) 式中任意的 $a_1, \cdots, a_s$, 设计的可卷偏差值都不会改变. 对于 C_8, 由表 2.1 可知, 最佳压缩比率非常接近 1, 因此从这种意义下, 可卷偏差的维数祸根问题并不严重.

2.3.5 混合偏差

虽然可卷偏差和中心化偏差可以克服 L_p-星偏差的许多缺点, 然而中心化偏差不满足准则 C_8, 而可卷偏差不满足 C_7. 因此, 需要寻找一个新的均匀性度量使其满足所有的 8 个准则 $C_1 - C_8$. 由于可卷偏差和中心化偏差都具有许多优良性质, 需要保留其优点并避免其缺点. 一个自然的想法是在一定意义下混合可卷偏差和中心化偏差这两个测度, 以得到一个新的均匀性度量.

从前面的讨论可知, 区域 $R_u(\boldsymbol{x})$ 的选择决定了相应的偏差. 因此, 需要给出一个更加合理的 $R_u(\boldsymbol{x})$. Zhou 等 (2013) 对中心化偏差和可卷偏差的 $R_u(\boldsymbol{x})$ 做一些修改并综合为一个新的 $R_u(\boldsymbol{x})$, 从而得到一个新的均匀性度量, 并称之为*混合偏差*.

可卷偏差中 $R_j^W(x_i, y_i)$ 的定义仅考虑 x_i 是否大于 y_i, 并没有考虑 x_i 和 y_i 之间的距离. 然而, 若 $R(\boldsymbol{x}_u)$ 定义在更大的区域上会更合理, 即在更大的区域上比较经验分布函数和均匀分布之间的差异. 因此, 可以对可卷偏差中的 $R_j^W(x_i, y_i)$ 做如下修改:

$$R_1^M(x_i, y_i) = \begin{cases} [\min(x_i, y_i), \max(x_i, y_i)], & |x_i - y_i| \geqslant \dfrac{1}{2}, \\ [0, \min(x_i, y_i)] \cup [\max(x_i, y_i), 1], & |x_i - y_i| < \dfrac{1}{2}, \end{cases}$$

$$R_1^M(\boldsymbol{x}_u, \boldsymbol{y}_u) = \bigotimes_{i \in u} R_1^M(x_i, y_i). \tag{2.26}$$

另外, 中心化偏差中的 $R_u^C(\boldsymbol{x}_u)$ 仅考虑 [0,1] 的小部分区域, 并忽略中心点, 这也是导致其具有维数祸根的原因. 因此, 对 $R_u^C(\boldsymbol{x}_u)$ 做如下修改：

$$R_2^M(x_i) = \begin{cases} [x_i, 1], & x_i \leqslant \dfrac{1}{2}, \\ [0, x_i], & x_i > \dfrac{1}{2}, \end{cases} \qquad R_2^M(\boldsymbol{x}_u) = \bigotimes_{i \in u} R_2^M(x_i), \tag{2.27}$$

其包含了每一维的中点. 图 2.7 和图 2.8 给出二维时 $R_1^M(\boldsymbol{x}, \boldsymbol{y})$ 和 $R_2^M(\boldsymbol{x})$ 的示意图, 其中 $R_1^M(\boldsymbol{x}, \boldsymbol{y})$ 的面积大于 1/4, 而在可卷偏差中 $R^W(\boldsymbol{x}, \boldsymbol{y})$ 的面积通常是小于 1/4 的. 在 $R_1^M(\boldsymbol{x}, \boldsymbol{y})$ 的定义中, C^2 中的每个点被取到的概率都是一样的; 而 $R_2^M(\boldsymbol{x})$ 通常包含中心点.

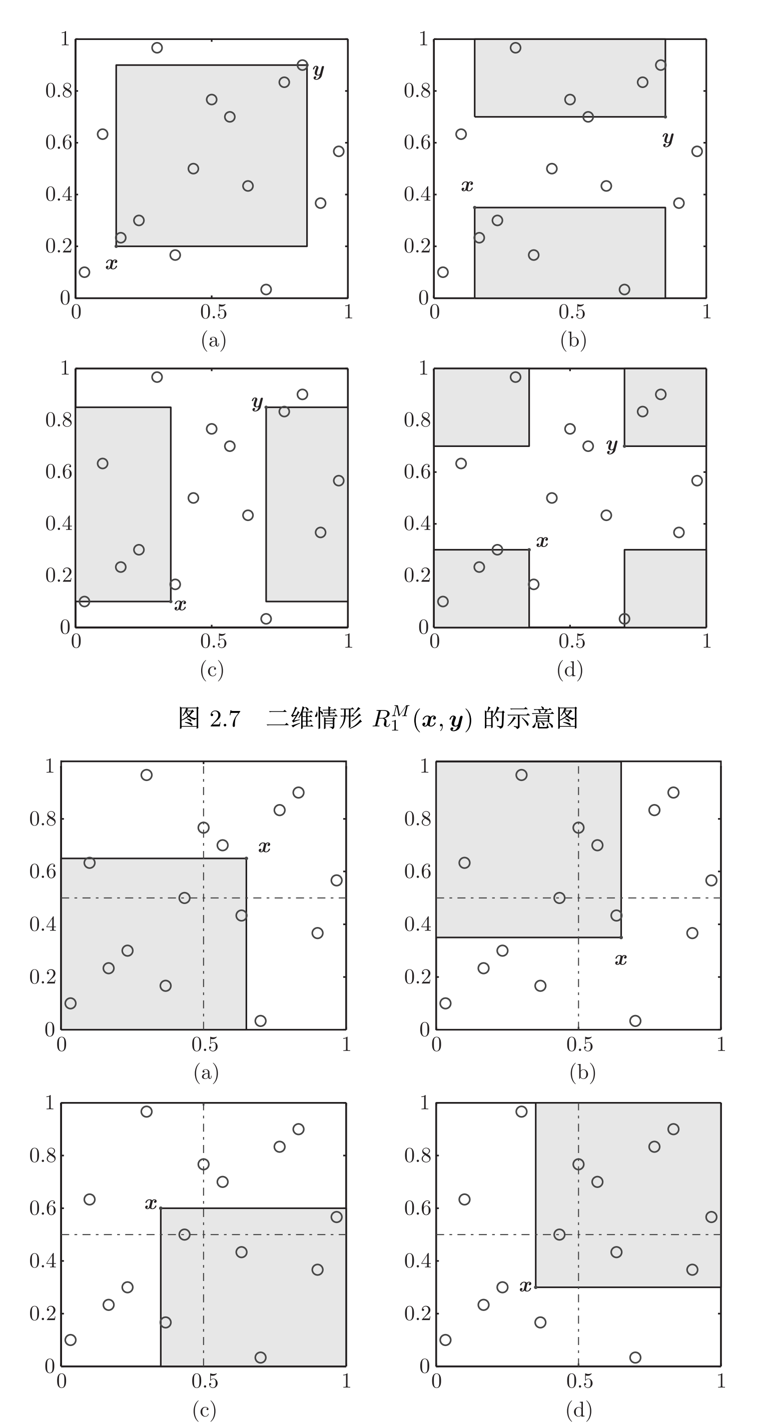


图 2.7　二维情形 $R_1^M(\boldsymbol{x},\boldsymbol{y})$ 的示意图

图 2.8　二维情形 $R_2^M(\boldsymbol{x})$ 的示意图

可以证明, 仅由 $R_1^M(\boldsymbol{x}_u,\boldsymbol{y}_u)$ 产生的偏差也具有类似于可卷偏差的特殊性质, 即对点集 $\mathcal{P}_0$ 中所有点都沿着方向 $(a_1,\cdots,a_s)$ 平移时, 偏差值不变. 类似地, $R_2^M(\boldsymbol{x})$ 更关注中心点附近的区域. 因此, 可以把这两个区域混合为一个新的区域 $R^M(x_j,y_j)$, 使其保留良好性质并克服不合理现象. 最简单的方法如下所示.

$$\begin{aligned} R^M(x_i,y_j) =& \frac{1}{2}R_1^M(x_i,y_j)+\frac{1}{2}R_2^M(x_i), \end{aligned} \tag{2.28}$$

$$\begin{aligned} R^M(\boldsymbol{x}_u,&\boldsymbol{y}_u) = \otimes_{i\in u}R^M(x_j,y_j),\\ \mathrm{disc}_u^M(\boldsymbol{x}_u,\boldsymbol{y}_u) =& \frac{1}{2}\left(\mathrm{Vol}(R_1^M(\boldsymbol{x}_u,\boldsymbol{y}_u))-\frac{\left|\mathcal{P}\cap R_1^M(\boldsymbol{x}_u,\boldsymbol{y}_u)\right|}{n}\right)\\ &+\frac{1}{2}\left(\mathrm{Vol}(R_2^M(\boldsymbol{x}_u))-\frac{\left|\mathcal{P}\cap R_2^M(\boldsymbol{x}_u)\right|}{n}\right), \end{aligned} \tag{2.29}$$

对于 C^s 上的点集 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$, Zhou 等 (2013) 给出混合偏差的表达式, 如下所示:

$$\begin{aligned} \mathrm{MD}^2(\mathcal{P}) =& \left(\frac{19}{12}\right)^s-\frac{2}{n}\sum_{i=1}^n\prod_{j=1}^s\left(\frac{5}{3}-\frac{1}{4}\left|x_{ij}-\frac{1}{2}\right|-\frac{1}{4}\left|x_{ij}-\frac{1}{2}\right|^2\right)\\ &+\frac{1}{n^2}\sum_{i=1}^n\sum_{k=1}^n\prod_{j=1}^s\left(\frac{15}{8}-\frac{1}{4}\left|x_{ij}-\frac{1}{2}\right|-\frac{1}{4}\left|x_{kj}-\frac{1}{2}\right|\right.\\ &\left.-\frac{3}{4}|x_{ij}-x_{kj}|+\frac{1}{2}|x_{ij}-x_{kj}|^2\right). \end{aligned} \tag{2.30}$$

易知, 混合偏差满足 C_1, C_3,C_4,C_5 和 C_7. 由于定义的区域 $R^M(\boldsymbol{x}_u,\boldsymbol{y}_u)$ 消除了原点的特殊地位, 并把所有的 2^s 个顶点同等看待, 则混合偏差满足 C_2. 例如, 从 (2.30) 可知, 反射变换 $x_i'=1-x_i$ 不改变混合偏差值. Zhou 等 (2013) 也证明了混合偏差也可以由再生核方法生成 (具体见 2.4 节), Hickernell (1998a) 证明由核方法生成的偏差满足 Koksma-Hlawka 不等式 (2.6), 从而也满足准则 C_6. 再考虑表 2.1 的例子, 可知混合偏差的最优的收缩比例 r 也非常接近 1, 故混合偏差的维数祸根现象不显著. 因此混合偏差满足所有的八个准则 C_1—C_8, 其可能比 L_p-星偏差、中心化偏差和可卷偏差更优. 表 2.2 列出常用的偏差如 L_p-星偏差、中心化偏差、可卷偏差和混合偏差的性质.

表 2.2　常用偏差的性质

准则	$L_p(p \neq 2)$	L_2	L_∞	CD	WD	MD
C_1	√	√	√	√	√	√
C_2	–	–	–	√	√	√
C_3	–	–	√	√	√	√
C_4	√	√	√	√	√	√
C_5	–	√	–	√	√	√
C_6	√	√	√	√	√	√
C_7	√	√	√	√	–	√
C_8	–	–	–	–	√	√

2.4　再　生　核

前几节中仅给出中心化偏差、可卷偏差和混合偏差的计算表达式, 而没有加以证明. 本节中, 我们介绍**再生核希尔伯特空间**这一工具, 并给出推导这些偏差的计算表达式的详细过程.

设 $\mathcal{X}$ 为 R^s 上的试验区域, $\mathcal{K}(\boldsymbol{z},\boldsymbol{t})$ 为定义在 $\mathcal{X}^2=\mathcal{X}\times\mathcal{X}$ 上并满足下面两条性质的实值函数:

(i) 对称性, 即

$$\mathcal{K}(\boldsymbol{z},\boldsymbol{t})=\mathcal{K}(\boldsymbol{t},\boldsymbol{z}),\quad \forall \boldsymbol{z},\boldsymbol{t}\in\mathcal{X};$$

(ii) 非负定性, 即

$$\sum_{i,k=1}^{n} c_i\mathcal{K}(\boldsymbol{z}_i,\boldsymbol{z}_k)c_k \geqslant 0,\quad \forall \boldsymbol{z}_i\in\mathcal{X},$$

其中 c_i 为实数, 称函数 $\mathcal{K}(\cdot,\cdot)$ 为**核函数**. 设 $\mathcal{W}$ 为 $\mathcal{X}$ 上与 $\mathcal{K}$ 相关的实值函数空间:

$$\mathcal{W}=\left\{F(\boldsymbol{x}):\int_{\mathcal{X}^2}\mathcal{K}(\boldsymbol{z},\boldsymbol{t})\mathrm{d}F(\boldsymbol{z})\mathrm{d}F(\boldsymbol{t})<\infty\right\}. \tag{2.31}$$

定义任意两个函数 $F,G\in\mathcal{W}$ 的内积为

$$\langle F,G\rangle_{\mathcal{W}}=\int_{\mathcal{X}^2}\mathcal{K}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}F(\boldsymbol{x})\mathrm{d}G(\boldsymbol{y}). \tag{2.32}$$

对任意函数 $F\in\mathcal{W}$, 定义其范数 $\|\cdot\|_{\mathcal{W}}$ 为 $\|F\|_{\mathcal{W}}=[\langle F,F\rangle_{\mathcal{W}}]^{1/2}$. 则 $\mathcal{W}$ 是 $\mathcal{X}$ 上的实值函数的希尔伯特空间, 且由 Cauchy-Schwarz 不等式可知, $\langle F,G\rangle_{\mathcal{W}}$ 值为有限数. 若核函数 $\mathcal{K}$ 满足下面的性质:

$$\begin{cases}\mathcal{K}(\cdot,\boldsymbol{x})\in\mathcal{W}, & \forall\ \boldsymbol{x}\in\mathcal{X},\\ F(x)=\langle\mathcal{K}(\cdot,\boldsymbol{x}),F\rangle_{\mathcal{W}}, & \forall\ F\in\mathcal{W},\ \boldsymbol{x}\in\mathcal{X},\end{cases} \tag{2.33}$$

则称该核函数为 $\mathcal{W}$ 的**再生核**, $\mathcal{W}$ 为再生核希尔伯特空间. Moore-Aronszajn 定理 (Aronszajn, 1950) 说明, 对任意给定的核函数 $\mathcal{K}$, 存在一个定义在 $\mathcal{X}$ 上的函数的希尔伯特空间, 且 $\mathcal{K}$ 恰好是其再生核.

应用再生核希尔伯特空间这一工具, 可以定义 L_2-偏差用于衡量设计的均匀性. 设 $\mathcal{X}$ 是一试验区域, F 是 $\mathcal{X}$ 上的均匀分布函数, 且 $F_{\mathcal{P}}$ 为 $\mathcal{X}$ 上的点集 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 的经验分布函数.

定义 2.3 衡量点集 $\mathcal{P}$ 均匀性的偏差定义为

$$D(\mathcal{P}, \mathcal{K}) = \|F - F_{\mathcal{P}}\|_{\mathcal{W}}, \tag{2.34}$$

其中 $\mathcal{W}$ 见 (2.31) 式.

(2.34) 式定义的偏差由核函数 $\mathcal{K}$ 确定, 其可以衡量 F 和 $F_{\mathcal{P}}$ 之间的差异. 一个核函数确定一种偏差. 对 (2.34) 式的积分式展开, 可得偏差的计算表达式:

$$\begin{aligned} D^2(\mathcal{P}, \mathcal{K}) =& \langle F - F_{\mathcal{P}}, F - F_{\mathcal{P}} \rangle_{\mathcal{W}} \\ =& \int_{\mathcal{X}^2} \mathcal{K}(\boldsymbol{t}, \boldsymbol{z}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \\ =& \int_{\mathcal{X}^2} \mathcal{K}(\boldsymbol{t}, \boldsymbol{z})\, \mathrm{d}F(\boldsymbol{t}) \mathrm{d}F(\boldsymbol{z}) - \frac{2}{n} \sum_{i=1}^{n} \int_{\mathcal{X}} \mathcal{K}(\boldsymbol{t}, \boldsymbol{x}_i)\, \mathrm{d}F(\boldsymbol{t}) \\ &+ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \mathcal{K}(\boldsymbol{x}_i, \boldsymbol{x}_k). \end{aligned} \tag{2.35}$$

当 s 个因子的水平值都可以独立地设定时, 试验区域 $\mathcal{X}$ 可取为 $[0,1]^u$. 此时, $\mathrm{d}F(\boldsymbol{t}) = \mathrm{d}\boldsymbol{t}$, 且与 $\mathcal{K}$ 相应的偏差如下所示.

$$\begin{aligned} D^2(\mathcal{P}, \mathcal{K}) =& \int_{[0,1]^{2s}} \mathcal{K}(\boldsymbol{t}, \boldsymbol{z})\, \mathrm{d}\boldsymbol{t} \mathrm{d}\boldsymbol{z} - \frac{2}{n} \sum_{i=1}^{n} \int_{[0,1]^s} \mathcal{K}(\boldsymbol{t}, \boldsymbol{x}_i)\, \mathrm{d}\boldsymbol{t} \\ &+ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \mathcal{K}(\boldsymbol{x}_i, \boldsymbol{x}_k). \end{aligned} \tag{2.36}$$

给定预先确定的区域 $R_u(\boldsymbol{x})$, 可以确定相应的再生核 $\mathcal{K}$, 从而确定相应的 L_2-偏差.

定理 2.1 (2.19) 式的广义 L_2-偏差可由 (2.34) 式确定, 且其核函数 $\mathcal{K}$ 可表示为

$$\mathcal{K}^R(\boldsymbol{t}, \boldsymbol{z}) = \sum_{u \subseteq \{1:s\}} \mathcal{K}_u^R(\boldsymbol{t}_u, \boldsymbol{z}_u), \tag{2.37}$$

其中

$$\mathcal{K}_u^R(\boldsymbol{t}_u, \boldsymbol{z}_u) = \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) 1_{R_u(\boldsymbol{x})}(\boldsymbol{z}_u) \, \mathrm{d}\boldsymbol{x}. \tag{2.38}$$

证明 由 (2.18), 局部投影偏差 $\mathrm{disc}_u^R(\boldsymbol{x}_u)$ 由 $R_u(\boldsymbol{x})$ 确定, 且可以重写为下面的积分形式:

$$\begin{aligned}
\mathrm{disc}_u^R(\boldsymbol{x}_u) &= \mathrm{Vol}(R_u(\boldsymbol{x})) - \frac{|\mathcal{P}_u \cap R_u(\boldsymbol{x})|}{n} \\
&= \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) \, \mathrm{d}F(\boldsymbol{t}) - \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) \, \mathrm{d}F_{\mathcal{P}}(\boldsymbol{t}) \\
&= \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}),
\end{aligned}$$

其中 F 和 $F_{\mathcal{P}}$ 分别为均匀分布函数和设计 $\mathcal{P}$ 的经验分布函数, $\boldsymbol{t} = (t_1, \cdots, t_s)$, $\boldsymbol{t}_u = (t_j)_{j \in u}$ 为 $\boldsymbol{t}$ 在 $[0,1]^u$ 上的投影. 这意味着局部投影偏差的平方 L_2-范数为

$$\begin{aligned}
\left\| \mathrm{disc}_u^R \right\|_2^2 &= \int_{[0,1]^u} \left| \mathrm{disc}_u^R(\boldsymbol{x}_u) \right|^2 \, \mathrm{d}\boldsymbol{x} \\
&= \int_{[0,1]^u} \int_{[0,1]^{2s}} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) 1_{R_u(\boldsymbol{x})}(\boldsymbol{z}_u) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{x} \\
&= \int_{[0,1]^{2s}} \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) 1_{R_u(\boldsymbol{x})}(\boldsymbol{z}_u) \mathrm{d}\boldsymbol{x} \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \\
&= \int_{[0,1]^{2s}} \mathcal{K}_u(\boldsymbol{t}_u, \boldsymbol{z}_u) \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}),
\end{aligned}$$

其中 $\mathcal{K}_u^R(\boldsymbol{t}_u, \boldsymbol{z}_u) = \int_{[0,1]^u} 1_{R_u(\boldsymbol{x})}(\boldsymbol{t}_u) 1_{R_u(\boldsymbol{x})}(\boldsymbol{z}_u) \, \mathrm{d}\boldsymbol{x}$, 如 (2.38) 所示. 则 (2.19) 式的广义 L_2-偏差可以重写为

$$\begin{aligned}
D_2^R(\mathcal{P}) &= \left\| \left(\left\| \mathrm{disc}_u^R \right\|_2 \right)_{u \subseteq \{1:s\}} \right\|_2 \\
&= \left\{ \sum_{u \subseteq \{1:s\}} \int_{[0,1]^{2s}} \mathcal{K}_u^R(\boldsymbol{t}_u, \boldsymbol{z}_u) \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \right\}^{1/2} \\
&= \left\{ \int_{[0,1]^{2s}} \sum_{u \subseteq \{1:s\}} \mathcal{K}_u^R(\boldsymbol{t}_u, \boldsymbol{z}_u) \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \right\}^{1/2} \\
&= \left\{ \int_{[0,1]^{2s}} \mathcal{K}^R(\boldsymbol{t}, \boldsymbol{z}) \, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \right\}^{1/2},
\end{aligned}$$

其中 $\mathcal{K}^R(\boldsymbol{t},\boldsymbol{z})=\sum_{u\subseteq\{1:s\}}\mathcal{K}_u^R(\boldsymbol{t}_u,\boldsymbol{z}_u)$, 如 (2.37) 所示. 因此 $R_u(\boldsymbol{x})$ 确定了 (2.38) 式中的函数 $\mathcal{K}_u^R(\boldsymbol{t}_u,\boldsymbol{z}_u)$, 其确定了广义 L_2-偏差的表达式. 定理得证.

给定预先确定的区域 $R_u(\boldsymbol{x})$, 为了计算可卷偏差、中心化偏差和混合偏差等广义 L_2-偏差的计算表达式, 可能需要对核函数 $\mathcal{K}^R(\boldsymbol{t},\boldsymbol{z})$ 做更多的限制. 通常假设核函数 $\mathcal{K}^R(\boldsymbol{t},\boldsymbol{z})$ 具有可分性, 此时称核函数为**可分核**, 即

$$\mathcal{K}(\boldsymbol{t},\boldsymbol{z})=\prod_{j=1}^{s}\mathcal{K}_j(t_j,z_j),\quad \forall\ \boldsymbol{t},\boldsymbol{z}\in\mathcal{X}. \tag{2.39}$$

当试验区域为每一边缘区域的笛卡儿积时, 即 $\mathcal{X}=\mathcal{X}_1\times\cdots\times\mathcal{X}_s$, 核函数可取为可分核. 实际上, (2.37) 式的函数 $\mathcal{K}^R(\boldsymbol{t},\boldsymbol{z})$ 是一个可分核, 因为区域 $R_u(\boldsymbol{x}_u)$ 为边缘区域的笛卡儿积. 因此核函数 $\mathcal{K}_u^R$ 也具有乘积形式

$$\mathcal{K}_u^R(\boldsymbol{t}_u,\boldsymbol{z}_u)=\prod_{j\in u}\tilde{\mathcal{K}}_1^R(t_j,z_j),$$

其中若一维区域 $R_1(x)$ 仅由一个点确定, 即

$$\tilde{\mathcal{K}}_1^R(t_j,z_j)=\int_0^1 1_{R_j(x)}(t_j)1_{R_j(x)}(z_j)\,\mathrm{d}x, \tag{2.40a}$$

若一维区域 $R_1(x)$ 由两个点确定, 即

$$\tilde{\mathcal{K}}_1^R(t_j,z_j)=\int_0^1\int_0^1 1_{R_j(x,y)}(t_j)1_{R_j(x,y)}(z_j)\,\mathrm{d}x\mathrm{d}y, \tag{2.40b}$$

则由二项式定理, 由式 (2.37) 定义的核函数 $\mathcal{K}^R$ 也具有乘积形式

$$\mathcal{K}^R(\boldsymbol{t},\boldsymbol{z})=\prod_{j=1}^{s}[1+\tilde{\mathcal{K}}_1^R(t_j,z_j)]. \tag{2.41}$$

因此, 给定区域 $R_u(\boldsymbol{x}_u)$ 或 $R_u(\boldsymbol{x}_u,\boldsymbol{y}_u)$, 可以从 (2.40) 获得 $\tilde{\mathcal{K}}_1^R(t_j,z_j)$, 把乘积形式 (2.41) 代入 (2.36), 可得相应偏差的显式表达式. 例如, 中心化偏差、可卷偏差和混合偏差的核函数分别如下所示.

(a)**中心化偏差**. 根据 2.3 节定义的 $R_u^C(\boldsymbol{x}_u)$, 相应的 $\tilde{\mathcal{K}}_1^C(t_j,z_j)$ 和 $\mathcal{K}^C(\boldsymbol{t},\boldsymbol{z})$ 如下所示:

$$\begin{aligned}\tilde{\mathcal{K}}_1^C(t_j,z_j)&=\int_0^1 1_{R_j^C(x)}(t_j)1_{R_j^C(x)}(z_j)\,\mathrm{d}x\\&=\begin{cases}1/2-\max(t_j,z_j), & t_j,z_j\leqslant 1/2,\\ \min(t_j,z_j)-1/2, & 1/2<t_j,z_j,\\ 0, & \text{否则}\end{cases}\\&=\frac{1}{2}\left|t_j-\frac{1}{2}\right|+\frac{1}{2}\left|z_j-\frac{1}{2}\right|-\frac{1}{2}\left|t_j-z_j\right|,\\ \mathcal{K}^C(\boldsymbol{t},\boldsymbol{z})&=\prod_{j=1}^{s}\left[1+\frac{1}{2}\left|t_j-\frac{1}{2}\right|+\frac{1}{2}\left|z_j-\frac{1}{2}\right|-\frac{1}{2}\left|t_j-z_j\right|\right].\end{aligned}\tag{2.42}$$

(b) **可卷偏差**. 相应的 $\tilde{\mathcal{K}}_1^W(t_j,z_j)$ 和 $\mathcal{K}^W(\boldsymbol{t},\boldsymbol{z})$ 如下所示:

$$\begin{aligned}\tilde{\mathcal{K}}_1^W(t_j,z_j)&=\int_0^1\int_0^1 1_{R_j^W(x,y)}(t_j)1_{R_j^W(x,y)}(z_j)\,\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2}-|t_j-z_j|+|t_j-z_j|^2,\\ \mathcal{K}^W(\boldsymbol{t},\boldsymbol{z})&=\prod_{j=1}^{s}\left[\frac{3}{2}-|t_j-z_j|+|t_j-z_j|^2\right].\end{aligned}\tag{2.43}$$

(c) **混合偏差**. 根据 (2.28), 设

$$\begin{aligned}k_j^1(z_j,t_j)&=\int_0^1\int_0^1 1_{R_1^M(x_j,y_j)}(z_j)1_{R_1^M(x_j,y_j)}(t_j)\mathrm{d}x_j\mathrm{d}y_j\\&=\frac{3}{4}-|z_j-t_j|+|z_j-t_j|^2,\\ k_j^2(z_j,t_j)&=\int_0^1 1_{R_2^M(x_j)}(z_j)1_{R_2^M(x_j)}(t_j)\mathrm{d}x_j\\&=1-\frac{1}{2}\left|z_j-\frac{1}{2}\right|-\frac{1}{2}\left|t_j-\frac{1}{2}\right|-\frac{1}{2}\left|z_j-t_j\right|,\end{aligned}$$

则 $\tilde{\mathcal{K}}_1^M(t,z)$ 和 $\mathcal{K}^M(\boldsymbol{t},\boldsymbol{z})$ 如下所示:

$$\begin{aligned}\tilde{\mathcal{K}}_1^M(t,z)&=\frac{1}{2}k_j^1(z_j,t_j)+\frac{1}{2}k_j^2(z_j,t_j)\\&=\frac{7}{8}-\frac{1}{4}\left|x_{ij}-\frac{1}{2}\right|-\frac{1}{4}\left|x_{kj}-\frac{1}{2}\right|-\frac{3}{4}\left|x_{ij}-x_{kj}\right|+\frac{1}{2}\left|x_{ij}-x_{kj}\right|^2,\end{aligned}$$

$$\mathcal{K}^M(\boldsymbol{t},\boldsymbol{z})=\prod_{j=1}^{s}\left[\frac{15}{8}-\frac{1}{4}\left|x_{ij}-\frac{1}{2}\right|-\frac{1}{4}\left|x_{kj}-\frac{1}{2}\right|-\frac{3}{4}\left|x_{ij}-x_{kj}\right|+\frac{1}{2}\left|x_{ij}-x_{kj}\right|^2\right].$$

得到这三种偏差的核函数后, 根据 (2.36), 可得这三种偏差的显式表达式分别如 (2.21), (2.23) 和 (2.30) 所示. Hickernell (1998a) 也给出了其他类型的偏差, 如全星偏差、对称偏差和不固定偏差等. 这些偏差也是由核函数的方法得到的.

需要指出的是, 中心化偏差、可卷偏差和混合偏差这三种偏差都是定义在单位立方体上. 在实际应用中, 可能考虑在其他试验区域 $\mathcal{X}$ 的偏差, 例如, 单纯形区域和超球体等. 设 $R(\boldsymbol{x})$ 为 $\mathcal{X}$ 上预先确定的区域, $\mathcal{P}$ 为 $\mathcal{X}$ 上的点集. 设 $\mathcal{X}$ 的体积有限, 即 $\mathrm{Vol}(\mathcal{X}) < \infty$. 则局部投影偏差为

$$\begin{aligned}\mathrm{disc}^R(\boldsymbol{x}) &= \frac{\mathrm{Vol}(R(\boldsymbol{x}))}{\mathrm{Vol}(\mathcal{X})} - \frac{|\mathcal{P} \cap R(\boldsymbol{x})|}{n} \\ &= \int_{\mathcal{X}} 1_{R(\boldsymbol{x})}(\boldsymbol{t})\, \mathrm{d}F(\boldsymbol{t}) - \int_{\mathcal{X}} 1_{R(\boldsymbol{x})}(\boldsymbol{t})\, \mathrm{d}F_{\mathcal{P}}(\boldsymbol{t}) \\ &= \int_{\mathcal{X}} 1_{R(\boldsymbol{x})}(\boldsymbol{t})\, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}), \end{aligned} \tag{2.44}$$

其中 F 为 $\mathcal{X}$ 上的均匀分布函数, $F_{\mathcal{P}}$ 为设计 $\mathcal{P}$ 的经验分布函数. 则局部投影偏差的平方 L_2-范数为

$$\begin{aligned}\left\|\mathrm{disc}^R\right\|_2^2 &= \int_{\mathcal{X}} \left|\mathrm{disc}^R(\boldsymbol{x})\right|^2 \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathcal{X}} \int_{\mathcal{X}^2} 1_{R(\boldsymbol{x})}(\boldsymbol{t}) 1_{R(\boldsymbol{x})}(\boldsymbol{z}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z})\, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathcal{X}^2} \int_{\mathcal{X}} 1_{R(\boldsymbol{x})}(\boldsymbol{t}) 1_{R(\boldsymbol{x})}(\boldsymbol{z}) \mathrm{d}\boldsymbol{x}\, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}) \\ &= \int_{\mathcal{X}^2} \mathcal{K}(\boldsymbol{t}, \boldsymbol{z})\, \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{t}) \mathrm{d}(F - F_{\mathcal{P}})(\boldsymbol{z}), \end{aligned} \tag{2.45}$$

其中函数 $\mathcal{K} : \mathcal{X}^2 \to \mathbb{R}$ 被定义为

$$\mathcal{K}(\boldsymbol{t}, \boldsymbol{z}) = \int_{\mathcal{X}} 1_{R(\boldsymbol{x})}(\boldsymbol{t}) 1_{R(\boldsymbol{x})}(\boldsymbol{z})\, \mathrm{d}\boldsymbol{x}. \tag{2.46}$$

则一般区域 $\mathcal{X}$ 上的广义 L_2-偏差可以如 (2.35) 所示. 通常情形下, 区域 $\mathcal{X}$ 不可分, 则核函数 $\mathcal{K}$ 也不可分.

2.5 水平数有限的偏差

在前面几节中, 我们都假设试验区域 $\mathcal{X}$ 是连通的, 比如单位立方体 $[0,1]^s$. 然而在很多实际问题中, 经常面临试验区域是离散点集的情形. 例如, 定性变量的水平数为有限数, 且为了试验方便, 定量变量也可能仅选择几个水平. 特别地, 在部分

因子设计中, 每个因子的水平数都是有限的, 并记为 $\tau_1, \cdots, \tau_{q_j}$. 因此, 一个 s 个因子的部分因子试验的试验区域可以写成下面的笛卡儿积的形式

$$\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_s, \quad 其中 \quad \mathcal{X}_j = \{\tau_1, \cdots, \tau_{q_j}\}. \tag{2.47}$$

当至少有两个因子的水平数是不同时, 称该设计为*非对称设计*; 当所有水平 $q_1 = \cdots = q_s = q$ 时, 称之为*对称设计*. 因此, 自然需要给出一些直接定义在有限点集的试验区域上的偏差. Hickernell 和 Liu (2002), Liu 和 Hickernell (2002) 与 Fang 等 (2003c) 考虑*离散偏差*, 其也被称为*分类偏差*; Zhou 等 (2008) 对水平数有限的设计提出 Lee *偏差*. 通常地, 离散偏差对二水平设计比较适用, 而 Lee 偏差可适用于多水平设计.

2.5.1 离散偏差

对于 n 次试验, s 个因子的部分因子设计 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$, 定义 $\delta_{ik}(\mathcal{P})$ 为第 i 行和第 j 行之间的*相遇数*:

$$\delta_{ik}(\mathcal{P}) = \delta(\boldsymbol{x}_i, \boldsymbol{x}_k) = \sum_{j=1}^{s} \delta_{x_{ij}x_{kj}}, \tag{2.48}$$

其中 δ_{tz} 为 Kronecker delta 函数, 即当 $t = z$ 时, $\delta_{tz} = 1$, 否则 $\delta_{tz} = 0$. 需要指出的是, $s - \delta_{ik}(\mathcal{P})$ 是第 i 行和第 k 行之间的 Hamming *距离*, 其定义为两行之间对应元素不同的个数.

根据 (2.36), 可以定义核函数并得到相应的偏差. 离散偏差的核函数 $\mathcal{K}$ 为

$$\mathcal{K}(\boldsymbol{t}, \boldsymbol{z}) = \prod_{j=1}^{s} \mathcal{K}_j(t_j, z_j), \tag{2.49}$$

其中

$$\mathcal{K}_j(t, z) = a^{\delta_{tz}} b^{1-\delta_{tz}} = \begin{cases} a, & t = z, \\ b, & t \neq z, \end{cases} \quad \frac{-a}{q-1} < b < a, \tag{2.50}$$

且条件 $\dfrac{-a}{q-1} < b < a$ 保证 $\mathcal{K}_j$ 的正定性. 因此, $\mathcal{K}(\boldsymbol{t}, \boldsymbol{z}) = a^{\delta(\boldsymbol{t},\boldsymbol{z})} b^{s-\delta(\boldsymbol{t},\boldsymbol{z})}$. 由 (2.36), 易得离散偏差的表达式如下

$$\mathrm{DD}^2(\mathcal{P}) = -\prod_{j=1}^{s} \left[\frac{a + (q_j - 1)b}{q_j}\right] + \frac{1}{n^2} \sum_{i,k=1}^{n} a^{\delta_{ik}(\mathcal{P})} b^{s-\delta_{ik}(\mathcal{P})}. \tag{2.51}$$

因此

$$\mathrm{DD}^2(\mathcal{P}) = -\prod_{j=1}^{s} \left[\frac{a + (q_j - 1)b}{q_j}\right] + \frac{a^s}{n} + \frac{2}{n^2} \sum_{1 \leqslant i < k \leqslant n} a^{\delta_{ik}(\mathcal{P})} b^{s-\delta_{ik}(\mathcal{P})}. \tag{2.52}$$

由 (2.52), 为了最小化离散偏差, 自然要求不同行之间的相遇数 $\delta_{ik}(\mathcal{P})$ 越小越好, 因为 $a > b$. 这意味着在离散偏差意义下的均匀设计把设计点尽量分散, 使不同设计点之间的差异越大越好.

对于对称设计, $q_1 = \cdots = q_s = q$, (2.52) 式中的离散偏差表达式变成

$$\mathrm{DD}^2(\mathcal{P}) = -\left[\frac{a+(q-1)b}{q}\right]^s + \frac{a^s}{n} + \frac{2b^s}{n^2}\sum_{1\leqslant i<k\leqslant n}\left(\frac{a}{b}\right)^{\delta_{ik}(\mathcal{P})}. \tag{2.53}$$

对于文献中常讨论的两水平设计 $\mathcal{P}$, 离散偏差变成下面形式

$$\mathrm{DD}^2(\mathcal{P}) = -\left(\frac{a+b}{2}\right)^s + \frac{a^s}{n} + \frac{2b^s}{n^2}\sum_{1\leqslant i<k\leqslant n}\left(\frac{a}{b}\right)^{\delta_{ik}(\mathcal{P})}. \tag{2.54}$$

由 (2.54), 易知两水平设计 $\mathcal{P}$ 的离散偏差可以表示为它们的 Hamming 距离的函数. 类似地, 也可以证明两水平设计的中心化偏差、可卷偏差和混合偏差也都可以表示为 Hamming 距离的函数, 其证明留为作业.

离散偏差并没有假设水平 $\tau_1, \tau_2, \cdots, \tau_j$ 的顺序, 因此 (2.51) 定义的偏差在水平的任意置换下都保持不变, 然而 C^s 上的中心化偏差、可卷偏差和混合偏差并不一定满足这些性质. 因此离散偏差也被称为分类偏差.

由于离散偏差是基于 Hamming 距离的, 其可能不适合构造多水平的均匀设计. 例如, 考虑下面的 U-型设计 $U(10;5^9)$

$$\begin{pmatrix}
1 & 5 & 5 & 3 & 4 & 3 & 3 & 1 & 1\\
5 & 1 & 4 & 3 & 1 & 2 & 5 & 2 & 2\\
5 & 3 & 2 & 1 & 4 & 4 & 2 & 4 & 4\\
3 & 5 & 1 & 2 & 5 & 2 & 2 & 3 & 3\\
3 & 2 & 5 & 4 & 3 & 1 & 1 & 2 & 4\\
4 & 3 & 4 & 5 & 2 & 5 & 1 & 3 & 1\\
4 & 4 & 3 & 1 & 5 & 1 & 4 & 1 & 2\\
2 & 2 & 3 & 5 & 1 & 4 & 3 & 5 & 3\\
2 & 4 & 1 & 4 & 2 & 3 & 5 & 4 & 5\\
1 & 1 & 2 & 2 & 3 & 5 & 4 & 5 & 5
\end{pmatrix},$$

其为离散偏差意义下的均匀设计, 参考 (Fang et al., 2002b). 然而, 该设计的可卷偏差值为 1.5442. 下面考虑交换该设计第一列中的 1 和 5, 其他都保持不动, 则变换后设计的可卷偏差值为 1.5417, 其比原设计的可卷偏差更小. 这个例子说明离散偏差不太适合用于构造多水平定量因子的均匀设计.

2.5.2　Lee 偏差

我们已指出 Hamming 距离并没有考虑水平数之间的绝对距离. 然而在编码理论中广泛应用的 Lee 距离可以克服这个缺点 (Roth, 2006). 对于 $x, y \in Z_q = \{1, \cdots, q\}$, x 和 y 之间的 Lee 距离为

$$\mathrm{Lee}(x, y) = \min(|x-y|, q-|x-y|).$$

例如, 当 $q=4$ 时, $\mathrm{Lee}(1,2) = \mathrm{Lee}(2,3) = \mathrm{Lee}(3,4) = \mathrm{Lee}(4,1) = 1$ 且 $\mathrm{Lee}(1,3) = \mathrm{Lee}(2,4) = 2$. 易知 Lee 距离包含某种可卷性.

从 Lee 距离出发, 可以定义如下的核函数

$$\mathcal{K}_j(t, z) = 1 - \min(|t-z|, 1-|t-z|), \tag{2.55a}$$

$$\mathcal{K}(\boldsymbol{t}, \boldsymbol{z}) = \prod_{j=1}^{s} \mathcal{K}_j(t_j, z_j), \quad \forall\ \boldsymbol{t}, \boldsymbol{z} \in \mathcal{X}. \tag{2.55b}$$

称由该核函数生成的偏差为 Lee *偏差*, 其可认为是离散偏差的推广. 下面给出对称设计和非对称设计的 Lee 偏差的计算表达式.

首先考虑对称设计. 设 $\mathcal{U}(n; q^s)$ 为所有 n 次试验, s 个因子, q 个水平的 U-型对称设计, 其每一列的元素来源于 q 个水平 $\{1, \cdots, q\}$ 且每个水平出现的次数一样多. 对任意设计 $\boldsymbol{U} = (u_{ij}) \in \mathcal{U}(n; q^s)$, 其试验区域为 $\{1, \cdots, q\}^s$, 即包含所有的 s 个因子的水平组合. 定义映射:

$$x_{ij} = \frac{2u_{ij}-1}{2q}, \quad i = 1, \cdots, n,\ j = 1, \cdots, s, \tag{2.56}$$

设计 $\boldsymbol{U}$ 的这 n 次试验变换至 $[0,1]^s$ 中的 n 个点, 记相应设计为 $\boldsymbol{X} = (x_{ij})$. 因此试验区域变为 $\mathcal{X} = \left\{\frac{1}{2q}, \cdots, \frac{2q-1}{2q}\right\}^s$. 为了方便, 定义

$$\alpha_{ij}^k = \min(|x_{ik} - x_{jk}|, 1 - |x_{ik} - x_{jk}|), \tag{2.57}$$

则对于奇数 q 和偶数 q 的平方 Lee 偏差分别如下所示:

$$\mathrm{LD}_{\mathrm{odd}}^2(\boldsymbol{U}) = \frac{1}{n} - \left(\frac{3}{4} + \frac{1}{4q^2}\right)^s + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \prod_{k=1}^{s} (1 - \alpha_{ij}^k), \tag{2.58a}$$

且

$$\mathrm{LD}_{\mathrm{even}}^2(\boldsymbol{U}) = \frac{1}{n} - \left(\frac{3}{4}\right)^s + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \prod_{k=1}^{s} (1 - \alpha_{ij}^k). \tag{2.58b}$$

注意设计 $\boldsymbol{U}$ 的 Lee 偏差是通过其导出矩阵 $\boldsymbol{X}$ 计算的. 定义 $\mathrm{LD}(\boldsymbol{U})=\mathrm{LD}(\boldsymbol{X})$.

下面考虑非对称 U-型设计 $U(n;q_1\times\cdots\times q_s)$ 的 Lee 偏差. 若某些 q_i 相等, 则记非对称设计为 $U(n;q_1^{s_1}\times\cdots\times q_t^{s_t})$, 其中 $s=\sum_{i=1}^t s_i$. 记 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 和 $\mathcal{U}(n;q_1^{s_1}\times\cdots\times q_t^{s_t})$ 分别为所有的 $U(n;q_1\times\cdots\times q_s)$ 和 $U(n;q_1^{s_1}\times\cdots\times q_t^{s_t})$. 不失一般性, 假设 $q_1,\cdots,q_r$ 为奇数, $q_{r+1},\cdots,q_t$ 为偶数, 其中 $0\leqslant r\leqslant t$. 当 $r=0$ 时, 意味着所有的水平都是偶数, 而 $r=t$ 意味着所有的水平都是奇数. 对任意因子 j, 由映射:

$$f:l\to(2l-1)/(2q_i),\quad l=1,\cdots,q_i, \tag{2.59}$$

则 $\boldsymbol{U}$ 的 n 次试验变换至 $[0,1]^s$ 中的 n 个点, 记相应设计为 $\boldsymbol{X}=(x_{ij})$. 则 $\boldsymbol{U}\in\mathcal{U}(n;q_1^{s_1},\cdots,q_t^{s_t})$ 的平方 Lee 偏差可表示为

$$\mathrm{LD}^2(\boldsymbol{U})=\frac{1}{n}-\left(\frac{3}{4}\right)^{\sum_{i=r+1}^t s_i}\prod_{i=1}^r\left(\frac{3}{4}+\frac{1}{4q_i^2}\right)^{s_i}+\frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^n\prod_{k=1}^s(1-\alpha_{ij}^k), \tag{2.60}$$

其中 α_{ij}^k 如 (2.57) 式定义.

由于 Lee 距离比 Hamming 距离更适合区分多水平中不同水平之间的距离, Lee 偏差也比离散偏差更适合于多水平设计.

2.6 几种偏差的下界

在 2.5 节里, 我们详细介绍了偏差在度量试验设计的均匀性方面起到的重要作用, 并介绍了几种常见的偏差. 众所周知, 偏差的下界不仅可以作为一个基准来搜索均匀设计, 而且还有助于证实一些好的设计是否是均匀的. 因此, 偏差下界的问题得到广泛讨论, 许多学者致力于研究各种偏差的下界. 对于给定的某一种偏差, 如果一个设计的偏差值达到了该偏差一个紧的下界, 则该设计关于此偏差是均匀的. 这里 “紧” 的意思是该下界在某些情形下是可以达到的. 在本节中, 我们将围绕偏差下界问题给出一些重要的结果. 为方便起见, 首先介绍两水平设计的相关结果, 然后再介绍高水平和混水平设计的有关结果. 本节主要涉及 CD, WD, MD, DD 和 LD 等几种常见的偏差.

为方便起见, 我们回顾一下有关符号和基本知识. 记 $D(n;q_1\times\cdots\times q_s)$ 为一个 $n\times s$ 的矩阵, 其第 j 列的元素为 $\{1,\cdots,q_j\}$. 当某些 q_j 相等时, 记为 $D(n;q_1^{r_1}\times\cdots\times q_m^{r_m})$, $r_1+\cdots+r_m=s$; 当所有的 q_j 都等于 q 时, 记为 $D(n;q^s)$. 记 $\mathcal{D}(n;q_1\times\cdots\times q_s)$ 为所有设计 $D(n;q_1\times\cdots\times q_s)$ 构成的集合. 类似地, 有记号 $\mathcal{D}(n;q_1^{r_1}\times\cdots\times q_m^{r_m})$ 和 $\mathcal{D}(n;q^s)$ 表示相应的集合. 在某些情形下, 下界是在 U-型约束下获得的, 因此在

U-型约束下, 设计集由 $\mathcal{D}(n;q_1\times\cdots\times q_s)$, $\mathcal{D}(n;q_1^{r_1}\times\cdots\times q_m^{r_m})$ 和 $\mathcal{D}(n;q^s)$ 相应地变为 $\mathcal{U}(n;q_1\times\cdots\times q_s)$, $\mathcal{U}(n;q_1^{r_1}\times\cdots\times q_m^{r_m})$ 和 $\mathcal{U}(n;q^s)$, 其定义具体见定义 1.6.

假定 $\mathcal{P}$ 是 $\mathcal{D}(n;q^s)$ 中的一个设计, 用 $\boldsymbol{x}=(x_1,\cdots,x_s)$ 来表示它的某个处理组合, 其中 $x_j\in\{0,\cdots,q-1\}$, $1\leqslant j\leqslant s$. 令$\mathcal{V}^*$ 为 $\mathcal{P}$ 的所有 $N\ (=q^s)$ 个处理组合按字典顺序排列所构成的集合. 对任意的 $\boldsymbol{x}\in\mathcal{V}^*$ 和 $\mathcal{P}\in\mathcal{D}(n;q^s)$, 记$\boldsymbol{y}_{\mathcal{P}}(x)$ 为处理组合 x 出现在设计 $\mathcal{P}$ 中的次数, $\boldsymbol{y}_{\mathcal{P}}$ 为以 $\boldsymbol{y}_{\mathcal{P}}(x)$ 为元素且按字典顺序排列所构成的 $N\times 1$ 向量.

令 $\boldsymbol{I}_q$ 和 $\boldsymbol{1}_q$ 分别表示 $q\times q$ 的单位阵和所有元素全为 1 的 $q\times 1$ 向量. 定义

$$\boldsymbol{L}(0)\ =\ \boldsymbol{1}_q^{\mathrm{T}},\quad \boldsymbol{L}(1)\ =\ \boldsymbol{I}_q,\quad \boldsymbol{J}_q=\boldsymbol{1}_q\boldsymbol{1}_q^{\mathrm{T}}. \tag{2.61}$$

$\boldsymbol{1}_q$, $\boldsymbol{I}_q$ 和 $\boldsymbol{J}_q$ 的 t 阶 Kronecker 乘积分别记为 $\boldsymbol{1}_q^{(t)}$, $\boldsymbol{I}_q^{(t)}$ 和 $\boldsymbol{J}_q^{(t)}$. 令 Ω 为所有长度为 q 的二元组的集合. 对任意的 $\boldsymbol{u}=(u_1,u_2,\cdots,u_s)\in\Omega$, 定义矩阵

$$\boldsymbol{G}(\boldsymbol{u})\ =\ \bigotimes_{j=1}^{s}\boldsymbol{L}(u_j). \tag{2.62}$$

注意到 $\boldsymbol{G}(\boldsymbol{x})$ 是一个 $q^{\sum_j u_j}\times N$ 的矩阵. 这里记号 $\otimes$ 表示 Kronecker 乘积.

任取 $\mathcal{P}\in\mathcal{D}(n;q^s)$ 的 m 列 $(\mathcal{P}_{l_1},\mathcal{P}_{l_2},\cdots,\mathcal{P}_{l_m})$, 定义

$$B_{l_1\cdots l_m}(\mathcal{P})=\sum_{\alpha_1,\cdots,\alpha_m}\left(n_{\alpha_1\cdots\alpha_m}^{(l_1\cdots l_m)}-n/q^m\right)^2,$$

其中 $n_{\alpha_1\cdots\alpha_m}^{(l_1\cdots l_m)}$ 是 $(\mathcal{P}_{l_1},\mathcal{P}_{l_2},\cdots,\mathcal{P}_{l_m})$ 的行取水平组合 $(\alpha_1,\cdots,\alpha_m)$ 的次数, 求和是对所有 q^m 个水平组合进行的. 若 $B_{l_1\cdots l_m}(d)=0$, 则以 $(d_{l_1},d_{l_2},\cdots,d_{l_m})$ 为列构成的子设计是一个强度为 m 的正交表. 此外, 对 $1\leqslant m\leqslant s$ 定义

$$B_m(\mathcal{P})=\sum_{1\leqslant l_1<\cdots<l_m\leqslant s}B_{l_1\cdots l_m}(\mathcal{P})\Big/\binom{s}{m}. \tag{2.63}$$

容易验证 $B_m(\mathcal{P})=0$ 当且仅当 $\mathcal{P}\in\mathcal{D}(n;q^s)$ 是强度为 m 的正交表. 因此, $B_m(\mathcal{P})$ 度量了设计 $\mathcal{P}$ 与强度为 m 的正交表的相似程度.

2.6.1 中心化 L_2-偏差的下界

Fang 和 Mukerjee(2000) 针对两水平正规正交设计首次给出了 CD 的下界, 随后沿着此方向, 许多相关的结果陆续获得. 对于正规的两水平 U-型设计, Fang 和 Mukerjee(2000) 给出了如下的下界.

定理 2.2 令 $\mathcal{P}$ 为 2^s 完全因子设计的部分设计, 包含 $n=2^{s-k}$ 次试验. 则 $\mathrm{CD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{CD}}^{(1)}$, 其中

$$\mathrm{LB}_{\mathrm{CD}}^{(1)}=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\sum_{r=0}^{s-k}\binom{s}{r}\frac{1}{8^r}+\frac{1}{n}\sum_{r=s-k+1}^{s}\frac{1}{4^r}.$$

证明 定义 $\boldsymbol{B}_0 = \dfrac{1}{4}\boldsymbol{I}_2 + \boldsymbol{J}_2$, $\boldsymbol{B}_s$ 为 $\boldsymbol{B}_0$ 的 s 阶 Kronecker 乘积. 容易验证 $\boldsymbol{B}_s = \sum_{u\in\Omega} \dfrac{1}{4^{\sum u_j}}\boldsymbol{G}(\boldsymbol{u})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})$. 因此,

$$\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{y}_{\mathcal{P}} = \sum_{u\in\Omega} \frac{1}{4^{\sum u_j}}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})\boldsymbol{y}_{\mathcal{P}}. \tag{2.64}$$

对于每个 $u\in\Omega$, 向量 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}$ 的维数为 $2^{\sum u_j}$; 而且 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}$ 的元素为整数且其和为 n. 因此, 若 $\sum u_j \leqslant s-k$, 则有 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})\boldsymbol{y}_{\mathcal{P}} \geqslant n^2/2^{\sum u_j}$; 若 $\sum u_j > s-k$, 则有 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}G(\boldsymbol{u})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})\boldsymbol{y}_{\mathcal{P}} \geqslant n$. 根据 (2.64) 有

$$\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{y}_{\mathcal{P}} \geqslant n^2\left[\sum_{r=0}^{s-k}\binom{s}{r}\frac{1}{8^r} + \frac{1}{n}\sum_{r=s-k+1}^{s}\frac{1}{4^r}\right]. \tag{2.65}$$

注意到

$$\mathrm{CD}^2(\mathcal{P}) = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{y}_{\mathcal{P}}. \tag{2.66}$$

由 (2.65) 式和 (2.66) 式可完成证明. 证毕.

注意到定理 2.2 中的下界可达当且仅当 $\mathcal{P}$ 中的 2^{s-k} 个试验构成一个强度为 $s-k$ 的二水平正交表. 由于这样的正交表常常不存在, 所以该下界通常是达不到的.

Fang 等 (2002c) 把定理 2.2 的结果推广到了二水平非正规设计, 并给出了 CD 的一个下界, 可同时适用于正规设计和非正规设计.

定理 2.3 令 $\mathcal{P}\in\mathcal{D}(n;2^s)$, 则 $\mathrm{CD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{CD}}^{(2)}$, 其中

$$\mathrm{LB}_{\mathrm{CD}}^{(2)} = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n^2}\sum_{r=0}^{s}\binom{s}{r}\frac{1}{4^r}[nf_r + z_r(f_r+1)], \tag{2.67}$$

f_r 是不超过 $n/2^r$ 的最大整数, $z_r = n - 2^r f_r$. 特别地, 若 $n = 2^{s-k}$ $(1\leqslant k\leqslant s-1)$, 则 $\mathrm{CD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{CD}}^{(1)}$.

证明 令 Ω_r 为 Ω 中长度为 s 的恰好有 r 个元素为 1 的二元组 u 构成的集合. 由 (2.64) 式有

$$\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{B}_s\boldsymbol{y}_{\mathcal{P}} = \sum_{r=0}^{s}\frac{1}{4^i}\left\{\sum_{u\in\Omega_r}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{u})\boldsymbol{y}_{\mathcal{P}}\right\}. \tag{2.68}$$

注意到对每个 $\boldsymbol{u}\in\Omega_r$, $2^r\times 1$ 向量 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}$ 的元素为非负整数且和为 n. 因此

$$\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}G(\boldsymbol{u})^{\mathrm{T}}G(\boldsymbol{u})\boldsymbol{y}_{\mathcal{P}} \geqslant f_r^2(2^r - z_r) + (f_r+1)^2 z_r = nf_r + z_r(f_r+1),$$

由 (2.66) 式和 (2.68) 式可得 $\mathrm{LB}_{\mathrm{CD}}^{(2)}$.

此外, 若 $n=2^{s-k}$, 则当 $r\leqslant s-k$ 时, $f_r=n/2^r$, $z_r=0$; 当 $r>s-k$ 时, $f_r=0$, $z_r=n$. 由 (2.67) 式可知 $\mathrm{CD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{CD}}^{(1)}$ 成立. 证毕.

为了改进定理 2.3 中的结果, Fang 等 (2003b) 发现 $\mathcal{P}$ 的 CD 值有如下两种不同的表达式：一是通过 $\mathcal{P}$ 的列的所有水平组合的分布来表达; 二是通过 $\mathcal{P}$ 的任意两行的 Hamming 距离来表达. 对于 U-型设计, 利用这两种表达式, Fang 等 (2003b) 获得了 CD 值的两个新下界, 我们把这两个下界陈述在下面的定理里.

定理 2.4　令$\mathcal{P}\in\mathcal{U}(n;2^s)$, 则 $\mathrm{CD}^2(\mathcal{P})\geqslant \max(\mathrm{LB}_{\mathrm{CD}}^{(c)},\mathrm{LB}_{\mathrm{CD}}^{(r)})$, 其中

$$\mathrm{LB}_{\mathrm{CD}}^{(c)}=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\left(\frac{9}{8}\right)^s+\frac{1}{n^2}\sum_{r=1}^{s}\frac{1}{4^r}\binom{s}{r}s_{n,r,2}\left(1-\frac{s_{n,r,2}}{2^r}\right),\tag{2.69}$$

$$\mathrm{LB}_{\mathrm{CD}}^{(r)}=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\frac{1}{n}\left(\frac{5}{4}\right)^s+\frac{n-1}{n}\left(\frac{5}{4}\right)^\lambda,\tag{2.70}$$

$s_{n,r,2}$ 为 n 除 2^r 的余数 ($n \bmod 2^r$), $\lambda=s(n-2)/[2(n-1)]$.

证明　由 Fang 等 (2003b) 有

$$\mathrm{CD}^2(\mathcal{P})=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\left(\frac{9}{8}\right)^s+\frac{1}{n^2}\sum_{r=1}^{s}\frac{1}{4^r}\binom{s}{r}B_r(\mathcal{P})\tag{2.71}$$

$$=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\frac{1}{n^2}\sum_{i=1}^{s}\sum_{j=1}^{s}\left(\frac{5}{4}\right)^{\delta_{ij}(\mathcal{P})}.\tag{2.72}$$

令 $n=t2^m+s_{n,r,2}$, 其中 t 为非负整数. 对于 $\mathcal{P}\in\mathcal{D}(n;q^s)$ 的每 r 个列 $(\mathcal{P}_{l_1},\mathcal{P}_{l_2},\cdots,\mathcal{P}_{l_r})$, 根据 $B_{l_1\cdots l_r}(\mathcal{P})$ 的定义可知, 当在 n 行中有 $s_{n,r,2}$ 个水平组合出现 $t+1$ 次, 其余的 $2^r-s_{n,r,2}$ 个水平组合出现 t 次时, $B_{l_1\cdots l_r}(\mathcal{P})$ 达到最小. 最小值为

$$s_{n,r,2}\left(1-\frac{s_{n,r,2}}{2^r}\right)^2+(2^r-s_{n,r,2})\left(-\frac{s_{n,r,2}}{2^r}\right)^2=s_{n,r,2}\left(1-\frac{s_{n,r,2}}{2^r}\right).$$

因此 $B_r(\mathcal{P})\geqslant s_{n,r,2}\left(1-\frac{s_{n,r,2}}{2^r}\right)$. 由 (2.71) 式可得 (2.69) 式中的下界 $\mathrm{LB}_{\mathrm{CD}}^{(c)}$.

为了获得下界 $\mathrm{LB}_{\mathrm{CD}}^{(r)}$, 定义集合 $\{\delta_{ij}(\mathcal{P}),1\leqslant i\neq j\leqslant n\}$ 上的均匀分布的随机变量 Y. 定义凸函数 $f(y)=\left(\frac{5}{4}\right)^y$, 则有 $E(Y)=\lambda$. 因此, 由 Jensen 不等式有

$$E[f(Y)]=\frac{1}{n(n-1)}\sum_{1\leqslant i\neq j\leqslant n}\left(\frac{5}{4}\right)^{\delta_{ij}(\mathcal{P})}\geqslant f[E(Y)]=\left(\frac{5}{4}\right)^\lambda.$$

注意到 $\delta_{ii}(\mathcal{P})=s$, 由 (2.72) 式可获得 (2.70) 式中的下界 $\mathrm{LB}_{\mathrm{CD}}^{(r)}$. 证毕.

为了达到 (2.70) 式中的下界, λ 必须为整数且 $\mathcal{P}$ 的任意两不同行之间的 Hamming 距离都相等. 然而, 许多二水平设计都达不到这些要求, 因此 Chatterjee 等 (2012b) 对于 U-型设计给出了下面改进的下界.

定理 2.5 设$\mathcal{P} \in \mathcal{U}(n;2^s)$, 则 $\mathrm{CD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{CD}}^{(3)}$, 其中

$$\mathrm{LB}_{\mathrm{CD}}^{(3)} = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{5}{4}\right)^s + \frac{1}{n^2}\left(\frac{5}{4}\right)^w \left(p + \frac{5}{4}q\right), \tag{2.73}$$

$p+q=n(n-1)$, $pw+q(w+1)=n(n-2)s/2$, w 是不超过 $s(n-2)/[2(n-1)]$ 的最大整数.

引理 2.6 (Chatterjee et al., 2012b) 对任意的 U-型设计 $\mathcal{P} \in \mathcal{U}(n;2^s)$ 和任意的整数 l 有

$$\sum_{i=1}^{n} \sum_{j(\neq i)=1}^{n} [\delta_{ij}(\mathcal{P})]^l \geqslant pw^l + q(w+1)^l. \tag{2.74}$$

定理 2.5 的证明 由 (2.72) 式有

$$\begin{aligned}
\mathrm{CD}^2(\mathcal{P}) &= \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{5}{4}\right)^s + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n} e^{\delta_{ij}(\mathcal{P})} \\
&= \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{5}{4}\right)^m + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\sum_{l=0}^{\infty} \frac{\alpha^l[\delta_{ij}(\mathcal{P})]^l}{l!},
\end{aligned}$$

其中 $\alpha = \ln(5/4)$. 由引理 2.6 可知

$$\begin{aligned}
\mathrm{CD}^2(\mathcal{P}) &\geqslant \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{4}{5}\right)^s + \frac{p}{n^2}\sum_{l=0}^{\infty}\frac{\alpha^l w^l}{l!} + \frac{q}{n^2}\sum_{l=0}^{\infty}\frac{\alpha^l(w+1)^l}{l!} \\
&= \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{5}{4}\right)^s + \frac{p}{n^2}e^{\alpha w} + \frac{q}{n^2}e^{\alpha(w+1)} \\
&= \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n}\left(\frac{5}{4}\right)^s + \frac{p}{n^2}\left(\frac{5}{4}\right)^w + \frac{q}{n^2}\left(\frac{5}{4}\right)^{w+1}.
\end{aligned}$$

证毕.

图 2.9 展示了 $\mathrm{LB}_{\mathrm{CD}}^{(c)}$, $\mathrm{LB}_{\mathrm{CD}}^{(r)}$ 与 $\mathrm{LB}_{\mathrm{CD}}^{(3)}$ 之间的差别. 注意到它们中没有一个下界对于所有的试验数 n 和因子数 s 的组合占绝对优势. 一般地, 对于较小的 n 与较大的 s, $\mathrm{LB}_{\mathrm{CD}}^{(3)}$ 要好于另两个下界, 即 $\mathrm{LB}_{\mathrm{CD}}^{(3)}$ 更适合于评价饱和或超饱和设计的均匀性. 对于较大的 n 与较小的 s, $\mathrm{LB}_{\mathrm{CD}}^{(c)}$ 要比另两个下界好.

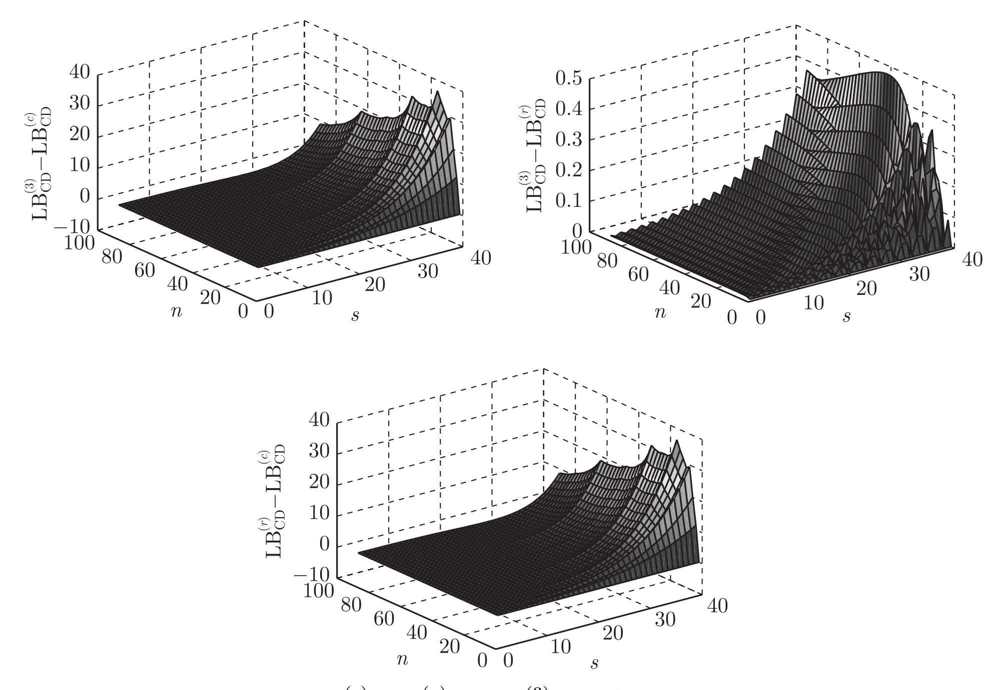


图 2.9　$\mathrm{LB}_{\mathrm{CD}}^{(c)}$, $\mathrm{LB}_{\mathrm{CD}}^{(r)}$ 与 $\mathrm{LB}_{\mathrm{CD}}^{(3)}$ 的比较 (文后附彩图)

对于高水平设计, 获得 CD 值的下界存在一些困难, 这是因为我们需要检验某个函数的根是否存在并且确保根是唯一的. Fang 等 (2006b) 分别对于 $\mathcal{P}\in\mathcal{D}(n;3^s)$ 和 $\mathcal{P}\in\mathcal{D}(n;4^s)$ 给出了 CD 值的下界. 对于 $\mathcal{P}\in\mathcal{D}(n;4^s)$, Elsawah 和 Qin (2014) 改进了 Fang 等 (2006b) 的结果并给出了 CD 值的一个更紧的下界. 对于混水平设计, Chatterjee 等 (2005) 首先研究了其 CD 值的下界问题, 并给出了一类混水平设计 $\mathcal{D}(n;p\times 2^s)$ 的 CD 值的下界, 其中 $p\geqslant 3$. 读者若有兴趣可查阅其相关文献.

2.6.2　可卷 L_2-偏差的下界

对于二水平设计, Fang 等 (2002c) 首先进行了研究并获得了 WD 如下的下界, 可适用于正规设计或非正规设计.

定理 2.7　设 $\mathcal{P}\in\mathcal{D}(n;2^s)$, 则 $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(21)}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(21)}=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\left(\frac{5}{4}\right)^s\sum_{r=0}^{s}\binom{s}{r}\frac{1}{5^r}[nf_r+z_r(f_r+1)],\tag{2.75}$$

f_r 和 z_r 如定理 2.3 所示. 特别地, 若 $n=2^{s-k}$ $(1\leqslant k\leqslant s-1)$, 则 $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(22)}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(22)}=-\left(\frac{4}{3}\right)^s+\left(\frac{5}{4}\right)^s\sum_{r=0}^{s-k}\binom{s}{r}\frac{1}{10^r}+\frac{1}{n}\left(\frac{5}{4}\right)^s\sum_{r=s-k+1}^{s}\binom{s}{r}\frac{1}{5^r}.$$

证明 注意到

$$\mathrm{WD}^2(\mathcal{P}) = -\left(\frac{4}{3}\right)^s + \frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{D}_s\boldsymbol{y}_{\mathcal{P}}, \tag{2.76}$$

其中 $\boldsymbol{D}_s$ 为 $\frac{1}{4}\boldsymbol{I}_2 + \frac{5}{4}\boldsymbol{J}_2$ 的 s 阶 Kronecker 乘积. 类似定理 2.3 的证明思路, 可获得下界 $\mathrm{LB}_{\mathrm{WD}}^{(21)}$. 证毕.

定理 2.7 中的下界通常是不可达的, 为了改进定理 2.7 中的下界, Fang 等 (2003b) 和 Chatterjee 等 (2012b) 对于 U-型设计分别给出了 WD 的新的下界, 我们总结在如下定理.

定理 2.8 设$\mathcal{P} \in \mathcal{U}(n; 2^s)$, 则

$$\mathrm{WD}^2(\mathcal{P}) \geqslant \max\left(\mathrm{LB}_{\mathrm{WD}}^{(2c)}, \mathrm{LB}_{\mathrm{WD}}^{(2r)}, \mathrm{LB}_{\mathrm{WD}}^{(23)}\right),$$

其中

$$\mathrm{LB}_{\mathrm{WD}}^{(2c)} = \left(\frac{11}{8}\right)^s - \left(\frac{4}{3}\right)^s + \frac{1}{n^2}\left(\frac{5}{4}\right)^s \sum_{r=1}^{s} \frac{1}{5^r}\binom{s}{r} s_{n,r,2}\left(1 - \frac{s_{n,r,2}}{2^r}\right),$$

$$\mathrm{LB}_{\mathrm{WD}}^{(2r)} = -\left(\frac{4}{3}\right)^s + \frac{1}{n}\left(\frac{3}{2}\right)^s + \frac{n-1}{n}\left(\frac{5}{4}\right)^s\left(\frac{6}{5}\right)^{\lambda} \tag{2.77}$$

和

$$\mathrm{LB}_{\mathrm{WD}}^{(23)} = -\left(\frac{4}{3}\right)^s + \frac{1}{n}\left(\frac{3}{2}\right)^s + \frac{1}{n^2}\left(\frac{5}{4}\right)^s\left[p\left(\frac{6}{5}\right)^w + q\left(\frac{6}{5}\right)^{w+1}\right]. \tag{2.78}$$

这里 $s_{n,r,2}$ 和 λ 如定理 2.4 所示, p, q 和 w 如定理 2.5 所示.

证明 由 (Fang et al., 2003b) 可知

$$\mathrm{WD}^2(\mathcal{P}) = \left(\frac{11}{8}\right)^s - \left(\frac{4}{3}\right)^s + \frac{1}{n^2}\sum_{r=1}^{s}\frac{1}{5^r}\binom{s}{r}B_r(\mathcal{P}) \tag{2.79}$$

$$= -\left(\frac{4}{3}\right)^s + \frac{1}{n^2}\left(\frac{5}{4}\right)^s \sum_{i=1}^{s}\sum_{j=1}^{s}\left(\frac{6}{5}\right)^{\delta_{ij}(\mathcal{P})}. \tag{2.80}$$

类似定理 2.4 和定理 2.5, 可获得下界 $\mathrm{LB}_{\mathrm{WD}}^{(2c)}$, $\mathrm{LB}_{\mathrm{WD}}^{(2r)}$ 及 $\mathrm{LB}_{\mathrm{WD}}^{(23)}$. 证毕.

图 2.10 表明: 对于较小的 n 与较大的 s, $\mathrm{LB}_{\mathrm{WD}}^{(23)}$ 要优于 $\mathrm{LB}_{\mathrm{WD}}^{(2c)}$ 和 $\mathrm{LB}_{\mathrm{WD}}^{(2r)}$, 对于较大的 n 与较小的 s, $\mathrm{LB}_{\mathrm{WD}}^{(2c)}$ 要好一些.

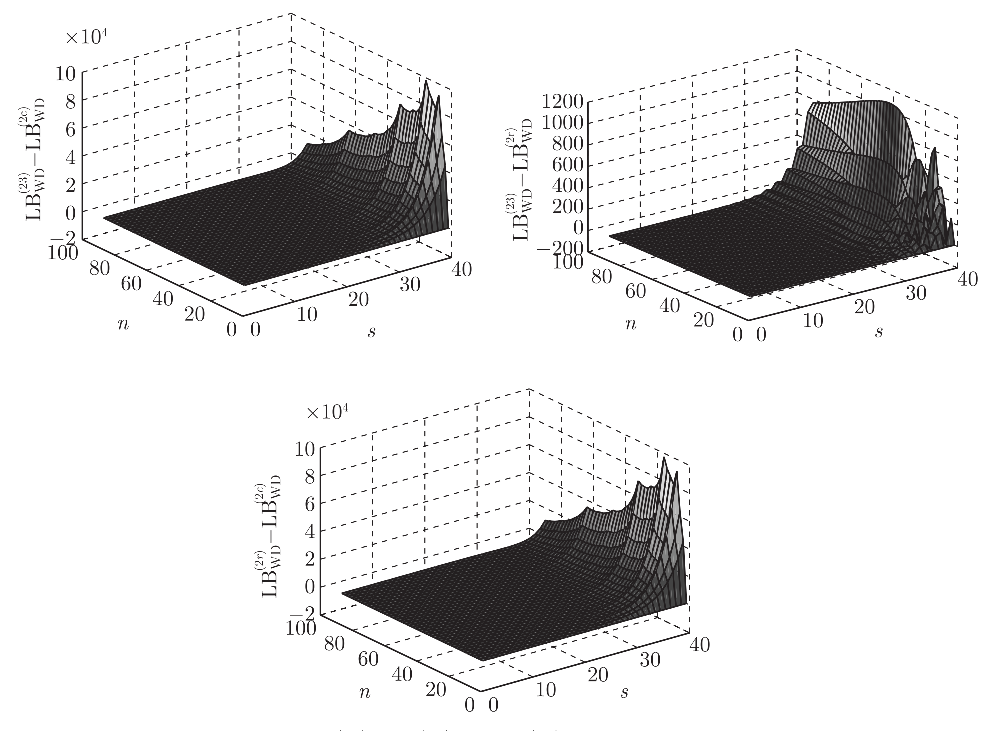


图 2.10　$\mathrm{LB}_{\mathrm{WD}}^{(2c)}$, $\mathrm{LB}_{\mathrm{WD}}^{(2r)}$ 与 $\mathrm{LB}_{\mathrm{WD}}^{(23)}$ 的比较 (文后附彩图)

对于三水平设计, Fang 等 (2002c) 给出了 WD 如下的下界.

定理 2.9　设 $\mathcal{P}\in\mathcal{D}(n;3^s)$, 则 $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(31)}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(31)}=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\left(\frac{23}{18}\right)^s\sum_{r=0}^{s}\binom{s}{r}\left(\frac{4}{23}\right)^r[ng_r+z_r(g_r+1)], \tag{2.81}$$

g_r 是不超过 $n/3^r$ 的最大整数, $t_r=n-3^rf_r$. 特别地, 若 $n=3^{s-k}$ $(1\leqslant k\leqslant s-1)$, 则 $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(32)}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(32)}=-\left(\frac{4}{3}\right)^s+\left(\frac{23}{18}\right)^s\sum_{r=0}^{s-k}\binom{s}{r}\left(\frac{4}{69}\right)^r+\frac{1}{n}\left(\frac{23}{18}\right)^s\sum_{r=s-k+1}^{s}\binom{s}{r}\left(\frac{4}{23}\right)^r.$$

证明　注意到

$$\mathrm{WD}^2(\mathcal{P})=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}Q_s\boldsymbol{y}_{\mathcal{P}}, \tag{2.82}$$

其中 Q_s 是 $\dfrac{2}{9}\boldsymbol{I}_3+\dfrac{23}{18}\boldsymbol{J}_3$ 的 s 阶 Kronecker 乘积. 类似定理 2.3 的证明, 可完成该定理的证明. 证毕.

类似地, 定理 2.9 中的下界通常是不可达的, 为了改进定理 2.9 中的下界, Fang 等 (2003b) 和 Zhang 等 (2015) 分别对于 U-型设计给出了 WD 新的下界, 总结在如下的定理中.

定理 2.10 设 $\mathcal{P} \in \mathcal{U}(n; 3^s)$ 和 $\Delta_3 = -\left(\frac{4}{3}\right)^s + \frac{1}{n}\left(\frac{3}{2}\right)^s$, 则

$$\mathrm{WD}^2(\mathcal{P}) \geqslant \max\left(\mathrm{LB}_{\mathrm{WD}}^{(3c)}, \mathrm{LB}_{\mathrm{WD}}^{(3r)}, \mathrm{LB}_{\mathrm{WD}}^{(33)}\right),$$

其中

$$\mathrm{LB}_{\mathrm{WD}}^{(3c)} = -\left(\frac{4}{3}\right)^s + \left(\frac{73}{54}\right)^s + \frac{1}{n^2}\left(\frac{23}{18}\right)^s \sum_{r=1}^{s}\left(\frac{4}{23}\right)^r \binom{s}{r} s_{n,r,3}\left(1 - \frac{s_{n,r,3}}{3^r}\right),$$

$$\mathrm{LB}_{\mathrm{WD}}^{(3r)} = \Delta_3 + \frac{n-1}{n}\left(\frac{23}{18}\right)^s \left(\frac{27}{23}\right)^{\lambda} \tag{2.83}$$

和

$$\mathrm{LB}_{\mathrm{WD}}^{(33)} = \Delta_3 + \frac{1}{n^2}\left(\frac{23}{18}\right)^s \left[p\left(\frac{27}{23}\right)^w + q\left(\frac{27}{23}\right)^{w+1}\right], \tag{2.84}$$

这里 $s_{n,r,3}$ 是 n 除 3^r 的余数 ($n \bmod 3^r$), $\lambda = s(n-3)/[3(n-1)]$, $p+q = n(n-1)$, $pw + q(w+1) = n(n-3)s/3$, w 为不超过 λ 的最大整数.

证明 由 (Fang et al., 2003b) 可知

$$\mathrm{WD}^2(\mathcal{P}) = -\left(\frac{4}{3}\right)^s + \left(\frac{73}{54}\right)^s + \frac{1}{n^2}\left(\frac{23}{18}\right)^s \sum_{r=1}^{s}\left(\frac{4}{23}\right)^r \binom{s}{r} B_r(\mathcal{P}) \tag{2.85}$$

$$= -\left(\frac{4}{3}\right)^s + \frac{1}{n^2}\left(\frac{23}{18}\right)^s \sum_{i=1}^{s}\sum_{j=1}^{s}\left(\frac{27}{23}\right)^{\delta_{ij}(\mathcal{P})}. \tag{2.86}$$

类似定理 2.4 和定理 2.5 的证明, 可获得下界 $\mathrm{LB}_{\mathrm{WD}}^{(3c)}$, $\mathrm{LB}_{\mathrm{WD}}^{(3r)}$ 及 $\mathrm{LB}_{\mathrm{WD}}^{(33)}$. 证毕.

图 2.11 表明: 对于较小的 n 和较大的 s, $\mathrm{LB}_{\mathrm{WD}}^{(33)}$ 要优于 $\mathrm{LB}_{\mathrm{WD}}^{(3c)}$ 和 $\mathrm{LB}_{\mathrm{WD}}^{(3r)}$, 对于较大的 n 和较小的 s, $\mathrm{LB}_{\mathrm{WD}}^{(3c)}$ 要好一些.

对于 q 水平设计, 其中 q 为正整数, Fang 等 (2005) 对于 U-型设计提出了 WD 的一个下界. 从方程 (2.23) 的解析表达式可知, $\mathrm{WD}^2(\mathcal{P})$ 是一个关于 $\alpha_{ij}^k \equiv |x_{il} - x_{jl}|(1 - |x_{il} - x_{jl}|)(i, j = 1, \cdots, n, i \neq j,\ k = 1, \cdots, s)$ 乘积的函数. 对于 U-型设计 $\mathcal{P} \in \mathcal{D}(n; q^s)$, 当 q 为偶数时, α-值只能取 $q/2+1$ 个可能值, 即 $0, 2(2q-2)/(4q^2)$, $4(2q-4)/(4q^2), \cdots, q^2/(4q^2)$; 当 q 为奇数时, α-值只能取 $(q+1)/2$ 个可能值, 即 0, $2(2q-2)/(4q^2)$, $4(2q-4)/(4q^2), \cdots, (q-1)(q+1)/(4q^2)$. 对于偶数和奇数 q, 表 2.3 给出了 α-值在集合 $\{\alpha_{ij}^k : 1 \leqslant i < j \leqslant n, 1 \leqslant k \leqslant s\}$ 上的分布.

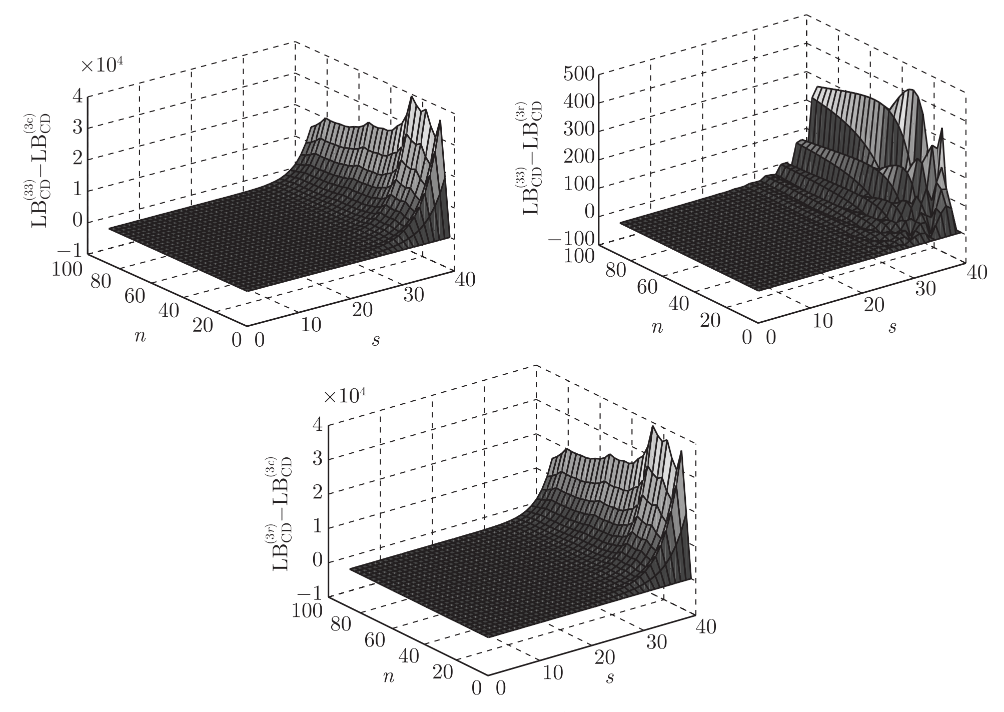


图 2.11　$\mathrm{LB}_{\mathrm{WD}}^{(3c)}$, $\mathrm{LB}_{\mathrm{WD}}^{(3r)}$ 与 $\mathrm{LB}_{\mathrm{WD}}^{(33)}$ 的比较 (文后附彩图)

表 2.3　设计 $\mathcal{P}\in\mathcal{D}(n;q^s)$ 的 α-值分布

q 偶数		q 奇数	
α-值	个数	α-值	个数
0	$\dfrac{sn(n-q)}{2q}$	0	$\dfrac{sn(n-q)}{2q}$
$\dfrac{2(2q-2)}{4q^2}$	$\dfrac{sn^2}{q}$	$\dfrac{2(2q-2)}{4q^2}$	$\dfrac{sn^2}{q}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\dfrac{(q-2)(q+2)}{4q^2}$	$\dfrac{sn^2}{q}$	$\dfrac{(q-3)(q+3)}{4q^2}$	$\dfrac{sn^2}{q}$
$\dfrac{q^2}{4q^2}$	$\dfrac{sn(n-q)}{2q}$	$\dfrac{(q-1)(q+1)}{4q^2}$	$\dfrac{sn(n-q)}{2q}$

定理 2.11　设 $\mathcal{P}\in\mathcal{U}(n;q^s)$, 则当 q 为偶数时, $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(\mathrm{even})}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(\mathrm{even})}=\Delta_3+\frac{n-1}{n}\left(\frac{3}{2}\right)^{\frac{s(n-q)}{q(n-1)}}\left(\frac{5}{4}\right)^{\frac{sn}{q(n-1)}}\prod_{i=1}^{q/2-1}\left[\frac{3}{2}-\frac{2i(2q-2i)}{4q^2}\right]^{\frac{2sn}{q(n-1)}};$$

当 q 为奇数时, $\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(\mathrm{odd})}$, 其中

$$\mathrm{LB}_{\mathrm{WD}}^{(\mathrm{odd})} = \Delta_3 + \frac{n-1}{n}\left(\frac{3}{2}\right)^{\frac{s(n-q)}{q(n-1)}}\left(\frac{5}{4}\right)^{\frac{sn}{q(n-1)}}\prod_{i=1}^{(q-1)/2}\left[\frac{3}{2}-\frac{2i(2q-2i)}{q^2}\right]^{\frac{2sn}{q(n-1)}}.$$

证明 由方程 (2.23) 可知, 对于给定的设计 $\mathcal{P}\in\mathcal{U}(n;q^s)$, 最小化 $\mathrm{WD}^2(\mathcal{P})$ 等价于关于 α_{ij}^k 最小化 $\sum_{1\leqslant i\neq j\leqslant n}\prod_{k=1}^s\left(\frac{3}{2}-\alpha_{ij}^k\right)$. 根据表 2.3 可知, 对给定的 (n,q,s), α-值的分布是相同的. 因此 $\prod_{1\leqslant i\neq j\leqslant n}\prod_{k=1}^s\left(\frac{3}{2}-\alpha_{ij}^k\right)$ 为常数且 $\frac{3}{2}-\alpha_{ij}^k>0$. 根据几何算术平均不等式, 对于 $1\leqslant i\neq j\leqslant n$, 当所有的 $\prod_{k=1}^s\left(\frac{3}{2}-\alpha_{ij}^k\right)$ 都相同时, $[\mathrm{WD}(d(\gamma))]^2$ 达到最小. $[\mathrm{WD}(d(\gamma))]^2$ 的下界表达式可由表 2.3 写出. 证毕.

当 $q=2,3$ 时, 定理 2.11 中的下界分别等价于 (2.77) 式中的 $\mathrm{LB}_{\mathrm{WD}}^{(2r)}$ 和 (2.83) 中的 $\mathrm{LB}_{\mathrm{WD}}^{(3r)}$.

对于混水平设计, Chatterjee 等 (2005) 首先研究了二、三混水平设计的 WD 的下界. 由于最优的二、三混水平设计在实际中应用广泛, 因此此类非对称设计的 WD 的下界显得非常重要.

定理 2.12 设 $\mathcal{P}\in\mathcal{D}(n;2^{s_1}\times 3^{s_2})$, 则

$$\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(1)},$$

其中

$$\mathrm{LB}_{\mathrm{WD}}^{(1)} = -\left(\frac{4}{3}\right)^{s_1+s_2}+\frac{1}{n^2}\left(\frac{5}{4}\right)^{s_1}\left(\frac{23}{18}\right)^{s_2}\sum_{i=0}^{s_1}\sum_{j=0}^{s_2}\binom{s_1}{i}\binom{s_2}{j}\left(\frac{1}{5}\right)^i\left(\frac{4}{23}\right)^j\theta_{ij}, \tag{2.87}$$

$\theta_{ij}=ng_{ij}+l_{ij}(g_{ij}+1)$, g_{ij} 是不超过 $n/(2^i3^j)$ 的最大整数, $l_{ij}=n-h_{ij}g_{ij}$, $0\leqslant i\leqslant s_1$, $0\leqslant j\leqslant s_2$.

定理 2.12 的证明类似于定理 2.3. 当 $s_1=0$ 或 $s_2=0$ 时, 定理 2.12 的下界等价于 (2.81) 式中的 $\mathrm{LB}_{\mathrm{WD}}^{(31)}$ 或 (2.75) 式的 $\mathrm{LB}_{\mathrm{WD}}^{(21)}$.

对于一般的 U-型混水平设计, Zhou 和 Ning (2008) 获得了设计 $\mathcal{P}\in\mathcal{U}(n;q_1^{s_1}\times\cdots\times q_m^{s_m})$ 的 WD 的下界. 不失一般性, 假定水平数 $q_1,\cdots,q_t$ 为奇数, 水平数 $q_{t+1},\cdots,q_m$ 为偶数, 其中 $0\leqslant t\leqslant s$, $t=0$ 意味着所有的水平数都为偶数, $t=m$ 意味着所有的水平数为奇数.

定理 2.13 设 $\mathcal{P}\in\mathcal{U}(n;q_1^{s_1}\times\cdots\times q_m^{s_m})$, $s=s_1+\cdots+s_m$ 以及 $\Delta_4=-\left(\frac{4}{3}\right)^s+\frac{1}{n}\left(\frac{3}{2}\right)^s$, 则

$$\mathrm{WD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{WD}}^{(3)},$$

其中

$$
\begin{aligned}
\mathrm{LB}_{\mathrm{WD}}^{(3)} =& \Delta_4 + \frac{n-1}{n}\left(\frac{3}{2}\right)^{\sum_{r=1}^{m}\frac{s_r(n-q_r)}{q_r(n-1)}}\left(\frac{5}{4}\right)^{\frac{n}{n-1}\sum_{r=t+1}^{m}\frac{s_r}{q_r}} \\
&\times \prod_{r=1}^{t}\prod_{i=1}^{(q_r-1)/2}\left[\frac{3}{2}-\frac{2i(2q_r-2i)}{4q_r^2}\right]^{\frac{2s_r n}{q_r(n-1)}} \\
&\times \prod_{r=t+1}^{m}\prod_{i=1}^{q_r/2-1}\left[\frac{3}{2}-\frac{2i(2q_r-2i)}{4q_r^2}\right]^{\frac{2s_r n}{q_r(n-1)}}.
\end{aligned}
$$

定理 2.13 的证明类似于定理 2.11. 当 $q_1 = \cdots = q_m = q$ 时, 定理 2.13 中的下界等价于定理 2.11 中的下界.

2.6.3　混合偏差的下界

对于二水平设计, Zhou 等 (2013) 给出了混合偏差如下定理的一个下界.

定理 2.14　设 $\mathcal{P} \in \mathcal{U}(n; 2^s)$, 则$\mathrm{MD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{MD}}^{(2:1)}$, 其中

$$
\mathrm{LB}_{\mathrm{MD}}^{(2:1)} = \left(\frac{19}{12}\right)^s - 2\left(\frac{305}{192}\right)^s + \left(\frac{39}{24}\right)^s + \frac{1}{n^2}\left(\frac{3}{2}\right)^s \sum_{r=1}^{s}\binom{s}{r}\frac{s_{n,r,2}}{6^r}\left(1-\frac{s_{n,r,2}}{2^r}\right)
$$

$s_{n,r,2}$ 如定理 2.4 所示.

证明　由 (Zhou et al., 2013) 可知

$$
\mathrm{MD}^2(\mathcal{P}) = \left(\frac{19}{12}\right)^s - 2\left(\frac{305}{192}\right)^s + \left(\frac{39}{24}\right)^s + \frac{1}{n^2}\left(\frac{3}{2}\right)^s \sum_{r=1}^{s}\frac{1}{6^r}\binom{s}{r}B_r(\mathcal{P}).
$$

类似定理 2.4 的证明, 可得下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$. 证毕.

为了达到下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$, $\mathcal{P}$ 要求为一个二水平正交设计. 显然, 在很多情形下该要求达不到. 而且, 即使当二水平设计 $\mathcal{P}$ 是正交的, 下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$ 也可能不是紧的. 因此, Ke 等 (2015) 给出了如下改进的下界.

定理 2.15　设 $\mathcal{P} \in \mathcal{U}(n; 2^s)$, λ 如定理 2.4 所示. 当 λ 是整数时, 则 $\mathrm{MD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{MD}}^{(2:2)}$, 其中

$$
\mathrm{LB}_{MD}^{(2:2)} = \left(\frac{19}{12}\right)^s - 2\left(\frac{305}{192}\right)^s + \frac{1}{n}\left(\frac{7}{4}\right)^s + \frac{n}{n-1}\left(\frac{3}{2}\right)^s\left(\frac{7}{6}\right)^{\lambda}.
$$

证明　由 Ke 等 (2015) 可知

$$
\mathrm{MD}^2(\mathcal{P}) = \left(\frac{19}{12}\right)^s - 2\left(\frac{305}{192}\right)^s + \left(\frac{39}{24}\right)^s + \frac{1}{n^2}\left(\frac{3}{2}\right)^s \sum_{i=1}^{s}\sum_{j=1}^{s}\left(\frac{7}{6}\right)^{\delta_{ij}(\mathcal{P})}. \quad (2.88)
$$

类似定理 2.4 的证明, 可得下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$. 证毕.

下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:2)}$ 可达当且仅当对所有的 $i \neq k$, $\delta_{ik}(\mathcal{P}) = \lambda$, 其中 $\delta_{ik}(\mathcal{P})$ 的定义见 (2.48) 式. 当 λ 不是整数时, 下界 $\mathrm{LB}_{\mathrm{MD}}^{(2:2)}$ 是不可达的. Elsawah 和 Qin(2015) 给出了如下下界, 该下界要优于 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$ 和 $\mathrm{LB}_{\mathrm{MD}}^{(2:2)}$.

定理 2.16 设 $\mathcal{P} \in \mathcal{U}(n; 2^s)$, 则 $\mathrm{MD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{MD}}^{(2:3)}$, 其中

$$\mathrm{LB}_{\mathrm{MD}}^{(2:3)} = \left(\frac{19}{12}\right)^s - 2\left(\frac{305}{192}\right)^s + \frac{1}{n}\left(\frac{7}{4}\right)^s + \frac{1}{n^2}\left(\frac{3}{2}\right)^s\left(\frac{7}{6}\right)^w\left(p + \frac{7}{6}q\right),$$

p, q 和 w 如定理 2.5 所示.

引理 2.17 (Elsawah and Qin, 2015) 假设 $\sum_{i=1}^n z_i = c$ 且 z_i 为非负数, 则对任意的整数 l 有

$$\sum_{i=1}^n l^{z_i} \geqslant l^w(\alpha + \beta l),$$

其中 α 和 β 为整数使得 $\alpha + \beta = n$, $\alpha w + \beta(w+1) = c$, w 为不超过 c/n 的最大整数.

定理 2.16 的证明 注意到对任意的设计 $\mathcal{P} \in \mathcal{D}(n; 2^s)$,

$$\sum_{i=1}^n \sum_{j(\neq i)=1}^n \delta_{ij}(\mathcal{P}) = ns(n-2)/2$$

为常数. 从引理 2.17 容易发现

$$\sum_{i=1}^s \sum_{j(\neq i)=1}^s \left(\frac{7}{6}\right)^{\delta_{ij}(\mathcal{P})} \geqslant \left(\frac{7}{6}\right)^w\left(p + \frac{7}{6}q\right). \tag{2.89}$$

联合 (2.88) 式和 (2.89) 式可完成证明. 证毕.

图 2.12 表明：对于较小的 n 和较大的 s, $\mathrm{LB}_{\mathrm{MD}}^{(2:3)}$ 要优于 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$, 对于较大的 n 和较小的 s, $\mathrm{LB}_{\mathrm{MD}}^{(2:3)}$ 要优于 $\mathrm{LB}_{\mathrm{MD}}^{(2:2)}$, 对于较小的 n 和较大的 s, $\mathrm{LB}_{\mathrm{MD}}^{(2:2)}$ 要优于 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}$.

对于三水平设计 $\mathcal{P} \in \mathcal{D}(n; 3^s)$, 令 φ 为不超过 $2s/3$ 的最大整数, $n_\varphi = (\varphi + 1)n - \dfrac{2ns}{3}$, $\varphi_0 = \dfrac{\ln\left(\dfrac{2}{3n}\right) + s\ln\left(\dfrac{9}{8}\right)}{\ln\left(\dfrac{42}{41}\right)}$. Ke 等 (2015) 对于 U-型设计 $\mathcal{P} \in \mathcal{U}(n; 3^s)$ 获得了 $\mathrm{MD}^2(\mathcal{P})$ 如下的下界.

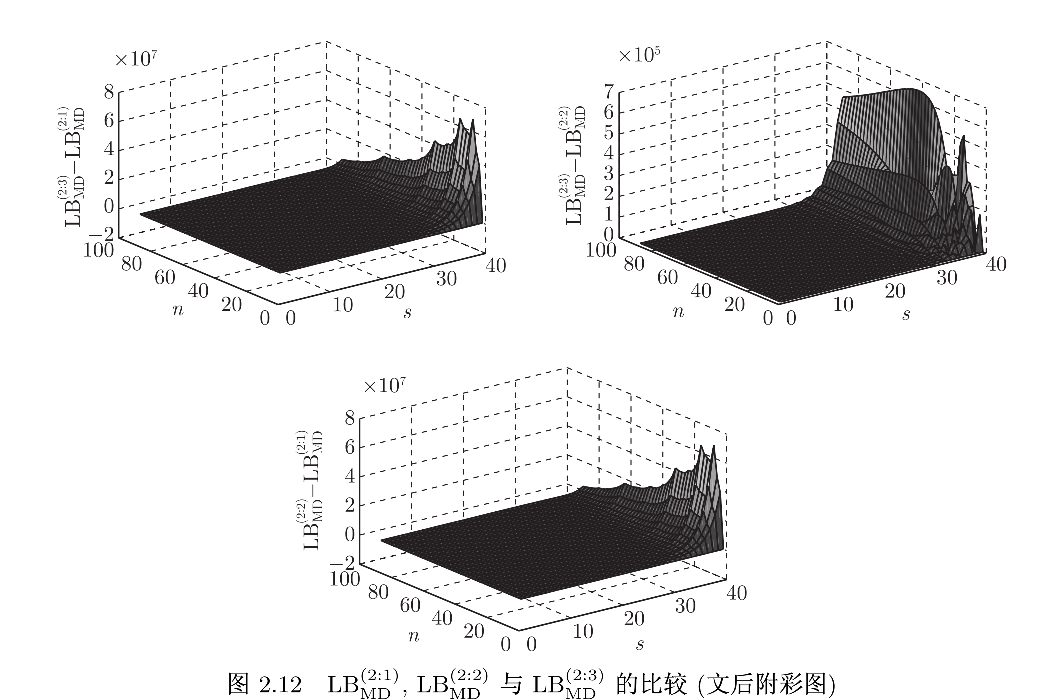


图 2.12 $\mathrm{LB}_{\mathrm{MD}}^{(2:1)}, \mathrm{LB}_{\mathrm{MD}}^{(2:2)}$ 与 $\mathrm{LB}_{\mathrm{MD}}^{(2:3)}$ 的比较 (文后附彩图)

定理 2.18 对任意的$\mathcal{P} \in \mathcal{U}(n; 3^s)$, 若 $\varphi_0 \geqslant \varphi$, 则 $\mathrm{MD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{MD}}^{(3:1)}$, 其中

$$\mathrm{LB}_{\mathrm{MD}}^{(3:1)} = \left(\frac{19}{12}\right)^s - \frac{2}{n}\left(\frac{5}{3}\right)^s \left[n_\varphi \left(\frac{14}{15}\right)^\varphi + (n-n_\varphi)\left(\frac{14}{15}\right)^{\varphi+1}\right]$$
$$+ \frac{1}{n^2}\left(\frac{15}{8}\right)^s \left[n_\varphi \left(\frac{41}{45}\right)^\varphi + (n-n_\varphi)\left(\frac{41}{45}\right)^{\varphi+1}\right] + \frac{n-1}{n}\left(\frac{15}{8}\right)^s e^{\widehat{\delta}}$$

以及

$$\widehat{\delta} = \frac{2s(n-3)}{9(n-1)}\ln\left(\frac{41}{45}\right) + \frac{4sn}{9(n-1)}\ln\left(\frac{23}{27}\right) + \frac{2sn}{9(n-1)}\ln\left(\frac{103}{135}\right).$$

证明 定义$\gamma_i = \left|\left\{i : x_{ik} \neq \frac{1}{2}\right\}\right|$, $\gamma_{ij}^{(1)} = \left|\left\{(i,j) : x_{ik} = x_{jk} = \frac{1}{6} \text{ 或 } \frac{5}{6}\right\}\right|$, $\gamma_{ij}^{(2)} = \left|\left\{(i,j) : (x_{ik}, x_{jk}) \in \left\{\left(\frac{1}{6}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{5}{6}\right)\right\}\right\}\right|$, $\gamma_{ij}^{(3)} = \left|\left\{(i,j) : (x_{ik}, x_{jk}) = \left(\frac{1}{6}, \frac{5}{6}\right)\right\}\right|$. 从 (2.30) 式可知

$$\mathrm{MD}^2(\mathcal{P}) = \left(\frac{19}{12}\right)^s - \frac{2}{n}\left(\frac{5}{3}\right)^s \sum_{i=1}^n \left(\frac{14}{15}\right)^{\gamma_i} + \frac{1}{n^2}\left(\frac{15}{8}\right)^s \sum_{i=1}^n \left(\frac{41}{45}\right)^{\gamma_i}$$
$$+ \frac{1}{n^2}\left(\frac{15}{8}\right)^s \sum_{i=1}^s \sum_{j(\neq i)=1}^s \left(\frac{41}{45}\right)^{\gamma_{ij}^{(1)}} \left(\frac{23}{27}\right)^{\gamma_{ij}^{(2)}} \left(\frac{103}{135}\right)^{\gamma_{ij}^{(3)}}. \tag{2.90}$$

定义 $g(\gamma_1,\cdots,\gamma_n)=-\dfrac{2}{n}\left(\dfrac{5}{3}\right)^s\sum_{i=1}^n\left(\dfrac{14}{15}\right)^{\gamma_i}+\dfrac{1}{n^2}\left(\dfrac{15}{8}\right)^s\sum_{i=1}^n\left(\dfrac{41}{45}\right)^{\gamma_i}$. 不难验证

$$\begin{aligned}g(\gamma_1,\cdots,\gamma_n)\geqslant&-\frac{2}{n}\left(\frac{5}{3}\right)^s\left[n_\varphi\left(\frac{14}{15}\right)^\varphi+(n-n_\varphi)\left(\frac{14}{15}\right)^{\varphi+1}\right]\\&+\frac{1}{n^2}\left(\frac{15}{8}\right)^s\left[n_\varphi\left(\frac{41}{45}\right)^\varphi+(n-n_\varphi)\left(\frac{41}{45}\right)^{\varphi+1}\right].\end{aligned}\tag{2.91}$$

若 $\varphi_0\geqslant\varphi$, 则下界 $g(\gamma_1,\cdots,\gamma_n)$ 是可达的.

类似 (2.74) 式中 $\mathrm{LB}_{\mathrm{CD}}^{(r)}$ 的证明有

$$\sum_{i=1}^s\sum_{j(\neq i)=1}^s\left(\frac{41}{45}\right)^{\gamma_{ij}^{(1)}}\left(\frac{23}{27}\right)^{\gamma_{ij}^{(2)}}\left(\frac{103}{135}\right)^{\gamma_{ij}^{(3)}}\geqslant n(n-1)e^{\widehat{\delta}}.\tag{2.92}$$

联合 (2.90) 式 –(2.92) 式可完成证明.

对任意的设计 $\mathcal{P}\in\mathcal{D}(n;3^s)$, 定义 $f_3(x)=\dfrac{4}{41}\left(\dfrac{41}{24}\right)^s\left(\dfrac{45}{41}\right)^x-\dfrac{n}{7}\left(\dfrac{14}{9}\right)^s\left(\dfrac{15}{14}\right)^x$, $n=3t$, $a_1=\ln\left(\dfrac{15}{8}\right)+2\ln\left(\dfrac{41}{24}\right)$, $a_2=\ln\left(\dfrac{103}{72}\right)+2\ln\left(\dfrac{115}{72}\right)$, $\zeta_1=st(t-1)$, $\zeta_2=2st^2$, ψ 是不超过 $s/3$ 的最大整数, $p_3+q_3=n$, $p_3\psi+q_3(\psi+1)=st$ 以及 $\zeta=a_1\zeta_1+a_2\zeta_2$. Elsawah 和 Qin (2015) 给出了 U-型设计 $\mathcal{P}\in\mathcal{U}(n;3^s)$ 的 $\mathrm{MD}^2(\mathcal{P})$ 的另一个下界.

定理 2.19 对任意的 $\mathcal{P}\in\mathcal{U}(n;3^s)$, 若 $f_3(\psi)\geqslant f_3(0)$, 则 $\mathrm{MD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{MD}}^{(3:2)}$, 其中

$$\begin{aligned}\mathrm{LB}_{\mathrm{MD}}^{(3:2)}=&\left(\frac{19}{12}\right)^s-\frac{2}{n}\left(\frac{14}{9}\right)^s\left(\frac{15}{14}\right)^\psi\left(p_3+\frac{15}{14}q_3\right)\\&+\frac{1}{n^2}\left(\frac{41}{24}\right)^s\left(\frac{45}{41}\right)^\psi\left(p_3+\frac{45}{41}q_3\right)+\frac{n-1}{n}e^{\frac{\zeta}{n(n-1)}}.\end{aligned}$$

证明 定义 $\varrho_{ij}=\left|\left\{(i,j):(x_{ik},x_{jk})\in\left\{\left(\frac{1}{6},\frac{1}{2}\right),\left(\frac{1}{2},\frac{5}{6}\right),\left(\frac{1}{2},\frac{1}{6}\right),\left(\frac{5}{6},\frac{1}{2}\right)\right\}\right\}\right|$, $\tau_{ij}=\left|\left\{(i,j):x_{ik}=x_{jk}\neq\frac{1}{2}\right\}\right|$, $\sigma_{ij}=\left|\left\{(i,j):(x_{ik},x_{jk})\in\left\{\left(\frac{1}{6},\frac{5}{6}\right),\left(\frac{5}{6},\frac{1}{6}\right)\right\}\right\}\right|$, $\lambda_{ij}=\left|\left\{(i,j):x_{ik}=x_{jk}=\frac{1}{2}\right\}\right|$. 由 (2.30) 式有

$$\begin{aligned}\mathrm{MD}^2(\mathcal{P})=&\left(\frac{19}{12}\right)^s-\frac{2}{n}\left(\frac{14}{9}\right)^s\sum_{i=1}^n\left(\frac{15}{14}\right)^{\lambda_{ii}}+\frac{1}{n^2}\left(\frac{41}{24}\right)^s\sum_{i=1}^n\left(\frac{45}{41}\right)^{\lambda_{ii}}\\&+\frac{1}{n^2}\sum_{i=1}^s\sum_{j(\neq i)=1}^s\left(\frac{15}{8}\right)^{\lambda_{ij}}\left(\frac{41}{24}\right)^{\tau_{ij}}\left(\frac{115}{72}\right)^{\varrho_{ij}}\left(\frac{103}{72}\right)^{\sigma_{ij}}.\end{aligned}\tag{2.93}$$

类似定理 2.18 的证明可完成该定理的证明. 证毕.

注意到, 定理 2.19 中的下界要优于定理 2.18 中的下界. 图 2.13 也表明：对于较大的 s, $\mathrm{LB}_{\mathrm{MD}}^{(3:2)}$ 要优于 $\mathrm{LB}_{\mathrm{MD}}^{(3:1)}$.

当 $s/3$ 为整数时, 即 $h=s/3$, 下界 $\mathrm{LB}_{\mathrm{MD}}^{(3:2)}$ 变为 $\left(\frac{19}{12}\right)^s-2\left(\frac{14}{9}\right)^s\left(\frac{15}{14}\right)^h+\frac{1}{n}\left(\frac{41}{24}\right)^s\left(\frac{45}{41}\right)^h+\frac{n-1}{n}e^{\frac{\zeta}{n(n-1)}}$.

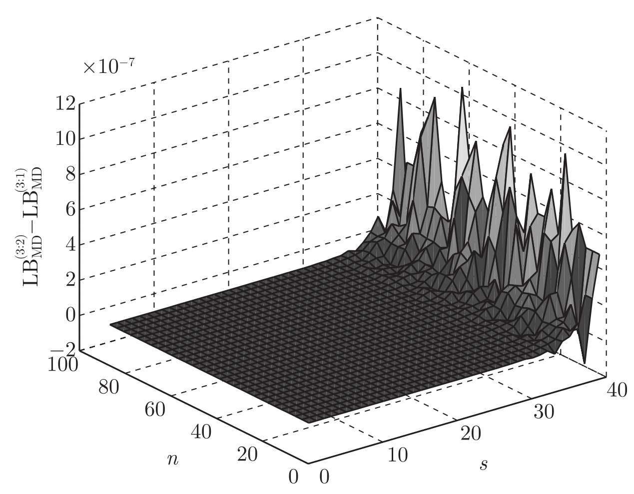


图 2.13　$\mathrm{LB}_{\mathrm{MD}}^{(3:1)}$ 与 $\mathrm{LB}_{\mathrm{MD}}^{(3:2)}$ 的比较 (文后附彩图)

类似地, 对于任意的设计 $\mathcal{P}\in\mathcal{D}(n;4^s)$, 定义 $f_4(x)=-\frac{144n}{1181}\left(\frac{1181}{768}\right)^s\left(\frac{1253}{1181}\right)^x+\frac{2}{27}\left(\frac{27}{16}\right)^s\left(\frac{29}{27}\right)^x$, $n=4\theta$, $b_1=\ln\left(\frac{53}{32}\right)+\ln\left(\frac{45}{32}\right)+2\ln\left(\frac{3}{2}\right)+2\ln\left(\frac{51}{32}\right)$, $b_2=\ln\left(\frac{27}{16}\right)+2\ln\left(\frac{29}{16}\right)$, $\eta_1=2s\theta^2$, $\eta_2=2s\theta(\theta-1)$, ω 是不超过 $s/2$ 的最大整数, $p_4+q_4=n$, $p_4\omega+q_4(\omega+1)=2s\theta$ 以及 $\eta=b_1\eta_1+b_2\eta_2$. Elsawah 和 Qin(2015) 对于 $\mathcal{P}\in\mathcal{U}(n;4^s)$ 给出了 $\mathrm{MD}^2(\mathcal{P})$ 如下的下界.

定理 2.20　*对任意的 U-型设计 $\mathcal{P}\in\mathcal{U}(n;4^s)$, 若 $f_4(\omega)\geqslant f_4(0)$, 则 $\mathrm{MD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{MD}}^{(4:1)}$, 其中*

$$\begin{aligned}\mathrm{LB}_{\mathrm{MD}}^{(4:1)}=&\left(\frac{19}{12}\right)^s-\frac{2}{n}\left(\frac{1181}{768}\right)^s\left(\frac{1253}{1181}\right)^\omega\left(p_4+\frac{1253}{1181}q_4\right)\\&+\frac{1}{n^2}\left(\frac{27}{16}\right)^s\left(\frac{29}{27}\right)^\omega\left(p_4+\frac{29}{27}q_4\right)+\frac{n-1}{n}e^{\frac{\eta}{n(n-1)}}.\end{aligned}$$

上述定理的证明类似于定理 2.19 和定理 2.20 的证明. 读者可查阅 (Elsawah and Qin, 2015) 来了解详细的证明过程.

特别地, 当 $s/2$ 是整数时, 即 $\omega=s/2$, 下界 $\mathrm{LB}_{\mathrm{MD}}^{(4:1)}$ 变为

$$\left(\frac{19}{12}\right)^s-2\left(\frac{1181}{768}\right)^s\left(\frac{1253}{1181}\right)^\omega+\frac{1}{n}\left(\frac{27}{16}\right)^s\left(\frac{29}{27}\right)^\omega+\frac{n-1}{n}e^{\frac{\eta}{n(n-1)}}.$$

对任意的非对称设计 $\mathcal{P}\in\mathcal{D}(n;2^{s_1}\times 3^{s_2})$, 设 $a_3=\ln\left(\frac{7}{6}\right)$, $\zeta_1^*=\frac{1}{9}s_2n(n-3)$, $\zeta_2^*=\frac{2}{9}s_2n^2$, $\zeta_3^*=\frac{1}{2}s_1n(n-2)$, μ 为不超过 $s_2/3$ 的最大整数, $p_5+q_5=n$, $p_5\mu+q_5(\mu+1)=\frac{ns_2}{3}$ 以及 $\zeta^*=a_1\zeta_1^*+a_2\zeta_2^*+a_3\zeta_3^*$, a_1 和 a_2 如定理 2.19 所示. 定义 $f_{23}(x)=\frac{4}{41}\left(\frac{7}{4}\right)^{s_1}\left(\frac{41}{24}\right)^{s_2}\left(\frac{45}{41}\right)^x-\frac{n}{7}\left(\frac{305}{192}\right)^{s_1}\left(\frac{14}{9}\right)^{s_2}\left(\frac{15}{14}\right)^x$. Elsawah 和 Qin (2016) 对于 $\mathcal{P}\in\mathcal{U}(n;2^{s_1}\times 3^{s_2})$ 获得了 $\mathrm{MD}^2(\mathcal{P})$ 如下的下界.

定理 2.21 对任意的 U-型设计 $\mathcal{P}\in\mathcal{U}(n;2^{s_1}\times 3^{s_2})$, 若 $f_{23}(\mu)\geqslant f_{23}(0)$, 则 $\mathrm{MD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{MD}}^{(23:1)}$, 其中

$$\begin{aligned}\mathrm{LB}_{\mathrm{MD}}^{(23:1)}=&\left(\frac{19}{12}\right)^{s_1+s_2}-\frac{2}{n}\left(\frac{305}{192}\right)^{s_1}\left(\frac{14}{9}\right)^{s_2}\left(\frac{15}{14}\right)^{\mu}\left(p_5+\frac{15}{14}q_5\right)\\&+\frac{1}{n^2}\left(\frac{7}{4}\right)^{s_1}\left(\frac{41}{24}\right)^{s_2}\left(\frac{45}{41}\right)^{\mu}\left(p_5+\frac{45}{41}q_5\right)+\frac{n-1}{n}\left(\frac{3}{2}\right)^{s_1}e^{\frac{\zeta^*}{n(n-1)}}.\end{aligned}$$

上述定理的证明类似于定理 2.19. 读者可查阅 (Elsawah and Qin, 2016).

当 $s_2/3$ 为整数, 即 $\mu=s_2/3$ 时, 则下界 $\mathrm{LB}_{\mathrm{MD}}^{(23:1)}$ 变为 $\left(\frac{19}{12}\right)^{s_1+s_2}-2\left(\frac{305}{192}\right)^{s_1}\cdot\left(\frac{14}{9}\right)^{s_2}\left(\frac{15}{14}\right)^{\mu}+\frac{1}{n}\left(\frac{7}{4}\right)^{s_1}\left(\frac{41}{24}\right)^{s_2}\left(\frac{45}{41}\right)^{\mu}+\frac{n-1}{n}\left(\frac{3}{2}\right)^{s_1}e^{\frac{\zeta^*}{n(n-1)}}$. 当 $s_1=0$ 或 $s_2=0$ 时, 定理 2.21 中的下界等价于定理 2.19 中的 $\mathrm{LB}_{\mathrm{MD}}^{(3:2)}$ 或定理 2.16 的 $\mathrm{LB}_{\mathrm{MD}}^{(2:3)}$.

2.6.4 离散偏差的下界

对于离散偏差, Fang 等 (2003c) 对于 U-型设计 $\mathcal{P}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$ 首先给出了 $\mathrm{DD}(\mathcal{P})$ 如下的下界.

定理 2.22 设 $\mathcal{P}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$, 则

$$\mathrm{DD}^2(\mathcal{P})\geqslant \mathrm{LB}_{\mathrm{DD}}^{(1)},$$

其中

$$\mathrm{LB}_{\mathrm{DD}}^{(1)}=-\prod_{j=1}^{s}\left[\frac{a+(q_j-1)b}{q_j}\right]+\frac{a^s}{n}+\frac{n-1}{n}b^s\left(\frac{a}{b}\right)^{\lambda},$$

且下界 $\mathrm{LB}_{\mathrm{DD}}^{(1)}$ 可达当且仅当 $\lambda=\sum_{j=1}^{s}(n/q_j-1)/(n-1)$ 是正整数且对所有的 $i\neq k$, $\delta_{ik}(\mathcal{P})=\lambda$.

证明　对于 $\mathcal{P} \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$ 容易验证: $\sum_{i=1}^n \sum_{j=1}^n \delta_{ij}(\mathcal{P}) = \sum_{i=1}^s \dfrac{n^2}{q_i}$, 即 $\sum_{1 \leqslant i<j \leqslant n} \delta_{ij}(\mathcal{P}) = \sum_{i=1}^s \dfrac{n^2}{q_i} - ns$ 为常数. 设 $\mathcal{P}^* \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$, 对于 $i \neq j$, $\delta_{ij}(\mathcal{P}^*) = \lambda$. 根据优势理论可知

$$(\delta_{12}(\mathcal{P}), \cdots, \delta_{1n}(\mathcal{P}), \delta_{23}(\mathcal{P}), \cdots, \delta_{2n}(\mathcal{P}), \cdots, \delta_{(n-1)n}(\mathcal{P}))\text{优势于}$$

$$(\delta_{12}(\mathcal{P}^*), \cdots, \delta_{1n}(\mathcal{P}^*), \delta_{23}(\mathcal{P}^*), \cdots, \delta_{2n}(\mathcal{P}^*), \cdots, \delta_{(n-1)n}(\mathcal{P}^*)).$$

注意到 (2.52) 式中的 $\mathrm{DD}^2(\mathcal{P})$ 是一个关于向量

$$(\delta_{12}(\mathcal{P}), \cdots, \delta_{1n}(\mathcal{P}), \delta_{23}(\mathcal{P}), \cdots, \delta_{2n}(\mathcal{P}), \cdots, \delta_{(n-1)n}(\mathcal{P}))$$

的 Schur 凹函数, 故有 $\mathrm{DD}^2(\mathcal{P}) \geqslant [\mathrm{DD}(\mathcal{P}^*)]^2$, 即 $\mathrm{DD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{DD}}^{(1)}$.

当 $\lambda = \sum_{j=1}^s (n/q_j - 1)/(n-1)$ 不是整数时, Qin 和 Fang(2004) 获得了 $\mathrm{LD}^2(\mathcal{P})$ 如下的下界并给出了设计 $\mathcal{P} \in \mathcal{U}(n; q_1 \times \cdots \times q_s)$ 达到该下界的充要条件. 证毕.

定理 2.23　设$\mathcal{P} \in \mathcal{U}(n; q_1 \times \cdots \times q_s)$, 则

$$\mathrm{DD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{DD}}^{(2)},$$

其中

$$\mathrm{LB}_{\mathrm{DD}}^{(2)} = -\prod_{j=1}^s \left[\frac{a + (q_j - 1)b}{q_j}\right] + \frac{a^s}{n} + \frac{(n-1)[b(\gamma + 1 - \lambda) + a(\lambda - \gamma)b^s}{nb} \left(\frac{a}{b}\right)^\gamma,$$

γ 是 λ 的整数部分. 下界 $\mathrm{LB}_{\mathrm{DD}}^{(2)}$ 可达当且仅当对 $\mathcal{P}$ 的任意第 i 行, $i \neq k$, $n-1$ 个 $\delta_{ik}(\mathcal{P})$ 值中, 有 $(n-1)(\gamma+1-\lambda)$ 个取 γ, $(n-1)(\lambda-\gamma)$ 个取 $\gamma+1$.

证明　设 $\mathcal{P}^* \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$, 对 $i \neq j$, 在 $n(n-1)$ 个 $\delta_{ij}(\mathcal{P}^*)$ 值, 有 $n(n-1)(\gamma+1-\lambda)$ 个取 γ, $n(n-1)(\lambda-\gamma)$ 个取 $\gamma+1$. 容易验证: $(\delta_{12}(\mathcal{P}), \cdots, \delta_{1n}(\mathcal{P}), \delta_{23}(\mathcal{P}), \cdots, \delta_{2n}(\mathcal{P}), \cdots, \delta_{(n-1)n}(\mathcal{P}))$ 被 $(\delta_{12}(\mathcal{P}^*), \cdots, \delta_{1n}(\mathcal{P}^*), \delta_{23}(\mathcal{P}^*), \cdots, \delta_{2n}(\mathcal{P}^*), \cdots, \delta_{(n-1)n}(\mathcal{P}^*))$ 优化. 根据定理 2.22 的证明可完成证明. 证毕.

注意到下界 $\mathrm{LB}_{\mathrm{DD}}^{(2)}$ 或 $\mathrm{LB}_{\mathrm{DD}}^{(1)}$ 是基于 $\mathcal{P}$ 的行的 Hamming 距离. 因此 $\mathrm{LB}_{\mathrm{DD}}^{(2)}$ 或 $\mathrm{LB}_{\mathrm{DD}}^{(1)}$ 对于评价近似饱和正交表或超饱和设计时非常有用.

当 $q_1 = \cdots = q_s = q$ 时, 由上面的定理可得如下结论.

定理 2.24　设$\mathcal{P} \in \mathcal{U}(n; q^s)$, 则

$$\mathrm{DD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{DD}}^{(3)},$$

其中

$$\mathrm{LB}_{\mathrm{DD}}^{(3)} = -\left[\frac{a + (q-1)b}{q}\right]^s + \frac{a^s}{n} + \frac{(n-1)[b(\gamma + 1 - \lambda) + a(\lambda - \gamma)b^s}{nb} \left(\frac{a}{b}\right)^\gamma,$$

$\lambda = s(n-q)/[q(n-1)]$, γ 是 λ 的整数部分.

基于 $\mathcal{P}$ 的列平衡性, Qin 和 Li (2006) 给出了 $\mathrm{DD}(\mathcal{P})$ 如下的下界.

定理 2.25 设 $\mathcal{P} \in \mathcal{D}(n; q^s)$, 则

$$\mathrm{DD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{DD}}^{(4)},$$

其中

$$\mathrm{LB}_{\mathrm{DD}}^{(4)} = \frac{b^s}{n^2} \sum_{i=1}^{s} \binom{s}{i} \left(\frac{a-b}{b}\right)^i R_{n,i,q} \left(1 - \frac{R_{n,i,q}}{q^i}\right),$$

$R_{n,i,q}$ 是 $n(\mathrm{mod}\ q^i)$ 的余数.

证明 注意到

$$\mathrm{DD}^2(\mathcal{P}) = \frac{b^s}{n^2} \sum_{i=1}^{s} \sum_{j=1}^{s} \binom{s}{i} \left(\frac{a-b}{b}\right)^i B_i(\mathcal{P}).$$

根据定理 2.4 的证明可完成证明. 证毕.

显然 $\mathrm{LB}_{\mathrm{DD}}^{(4)}$ 要优于 $\mathrm{LB}_{\mathrm{DD}}^{(3)}$, 这在评价正交表的均匀性非常有用.

对 $1 \leqslant j \leqslant s$, 设 ω_j 是不超过 n/q^j 的最大整数, $\theta_j = n - q^j \omega_j$ 以及 $\theta_j^* = n\omega_j + \theta_j(1+\omega_j)$. Chatterjee 和 Qin (2008) 对于 $\mathcal{P} \in \mathcal{D}(n; q^s)$ 给出了 $\mathrm{DD}(\mathcal{P})$ 如下的下界.

定理 2.26 设 $\mathcal{P} \in \mathcal{D}(n; q^s)$, 则

$$\mathrm{DD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{DD}}^{(5)},$$

其中

$$\mathrm{LB}_{\mathrm{DD}}^{(5)} = -\left[\frac{a+(q-1)b}{q}\right]^s + b^s + \frac{b^s}{n^2} \sum_{i=1}^{s} \binom{s}{i} \left(\frac{a-b}{b}\right)^i \theta_j^*.$$

证明 容易验证

$$\mathrm{DD}^2(\mathcal{P}) = -\left[\frac{a+(q-1)b}{q}\right]^s + \frac{1}{n^2} \boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}} M_s \boldsymbol{y}_{\mathcal{P}},$$

其中 M_s 是 $(a-b)\boldsymbol{I}_s + b\boldsymbol{J}_s$ 的 s 阶 Kronecker 乘积. 根据定理 2.3 的证明可完成证明.

根据 (Qin and Fang, 2004), 对任意的设计 $\mathcal{P} \in \mathcal{D}(n; q^s)$ 有下面的等式成立.

(I) 当 $q = 2$, $a = 5/4$, $b = 1$ 时, 有

$$\mathrm{DD}^2(\mathcal{P}) = \mathrm{CD}^2(\mathcal{P}) + 2\left(\frac{35}{32}\right)^s - \left(\frac{13}{12}\right)^s - \left(\frac{9}{8}\right)^s;$$

(II) 当 $q=2$, $a=3/2$, $b=5/4$ 时, 有

$$\mathrm{DD}^2(\mathcal{P})=\mathrm{WD}^2(\mathcal{P})+\left(\frac{4}{3}\right)^s-\left(\frac{11}{8}\right)^s;$$

(III) 当 $q=3$, $a=3/2$, $b=23/18$ 时, 有

$$\mathrm{DD}^2(\mathcal{P})=\mathrm{WD}^2(\mathcal{P})+\left(\frac{4}{3}\right)^s-\left(\frac{73}{54}\right)^s.$$

容易验证: 当 $q=2$ 时, 若 $a=5/4$, $b=1$, 则定理 2.24 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(3)}$、定理 2.25 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(4)}$ 和定理 2.26 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(5)}$ 分别为定理 2.4 中的下界 $\mathrm{LB}_{\mathrm{CD}}^{(r)}$、定理 2.4 中的下界 $\mathrm{LB}_{\mathrm{CD}}^{(c)}$ 和定理 2.3 中的下界 $\mathrm{LB}_{\mathrm{CD}}^{(2)}$; 若 $a=3/2$, $b=5/4$, 则定理 2.24 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(3)}$、定理 2.25 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(4)}$ 和定理 2.26 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(5)}$ 分别是定理 2.8 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(2r)}$、定理 2.8 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(2c)}$ 和定理 2.7 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(21)}$. 当 $q=3$ 时, 若 $a=3/2$, $b=23/18$, 则定理 2.24 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(3)}$、定理 2.25 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(4)}$ 和定理 2.26 中的下界 $\mathrm{LB}_{\mathrm{DD}}^{(5)}$ 分别等价于定理 2.10 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(3r)}$、定理 2.10 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(3c)}$ 和定理 2.9 中的下界 $\mathrm{LB}_{\mathrm{WD}}^{(31)}$.

2.6.5　Lee 偏差的下界

对于 q 水平 U-型设计, 其中 q 为正整数, Zhou 等 (2008) 首先给出了 LD 如下定理中的下界.

定理 2.27　设 $\mathcal{P}\in\mathcal{U}(n;q^s)$, 则当 q 为偶数时, $\mathrm{LD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{LD}}^{(\mathrm{even})}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(\mathrm{even})}=\frac{1}{n}-\left(\frac{3}{4}\right)^s+\frac{n-1}{n}\left(\frac{1}{2}\right)^{\frac{ns}{q(n-1)}}\left[\prod_{i=1}^{q/2-1}\frac{q+2i}{2q}\right]^{\frac{2sn}{q(n-1)}};$$

当 q 为奇数时, $\mathrm{LD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{LD}}^{(\mathrm{odd})}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(\mathrm{odd})}=\frac{1}{n}-\left(\frac{3}{4}+\frac{1}{4q^2}\right)^s+\frac{n-1}{n}\left[\prod_{i=1}^{(q-1)/2}\frac{q+2i-1}{2q}\right]^{\frac{2sn}{q(n-1)}}.$$

定理 2.27 的证明类似于定理 2.11.

当 $q=2,3$ 时, 定理 2.27 中的下界如下面的定理所示.

定理 2.28　设 $\mathcal{P}\in\mathcal{U}(n;q^s)$, 则当 $q=2$ 时, $\mathrm{LD}^2(\mathcal{P})\geqslant\mathrm{LB}_{\mathrm{LD}}^{(21)}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(21)}=\frac{1}{n}-\left(\frac{3}{4}\right)^s+\frac{n-1}{n}\left(\frac{1}{2}\right)^{\frac{ns}{2(n-1)}};$$

当 $q=3$ 时, $\mathrm{LD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{LD}}^{(31)}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(31)} = \frac{1}{n} - \left(\frac{7}{9}\right)^s + \frac{n-1}{n}\left(\frac{2}{3}\right)^{\frac{2sn}{3(n-1)}}.$$

当 $q=2,3$ 时, Zou 等 (2009) 给出了改进的下界, 要优于定理 2.28 中的下界.

定理 2.29　设 $\mathcal{P} \in \mathcal{D}(n;q^s)$ 为 U-型设计, 则当 $q=2$ 时, $\mathrm{LD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{LD}}^{(22)}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(22)} = -\left(\frac{3}{4}\right)^s + \frac{1}{2^s n^2}\sum_{r=0}^{s}\binom{s}{r}[nf_r + z_r(f_r+1)],$$

f_r 和 z_r 如定理 2.3 所示; 当 $q=3$ 时, $\mathrm{LD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{LD}}^{(32)}$, 其中

$$\mathrm{LB}_{\mathrm{LD}}^{(32)} = -\left(\frac{7}{9}\right)^s + \frac{1}{3^s n^2}\sum_{r=0}^{s}\binom{s}{r}2^{s-r}[ng_r + t_r(g_r+1)],$$

g_r 和 t_r 如定理 2.9 所示.

证明　当 $q=2$ 时, 有

$$\mathrm{LD}^2(\mathcal{P}) = -\left(\frac{3}{4}\right)^s + \frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}} L_s^{(2)} \boldsymbol{y}_{\mathcal{P}};$$

当 $q=3$ 时, 有

$$\mathrm{LD}^2(\mathcal{P}) = -\left(\frac{7}{9}\right)^s + \frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}} L_s^{(3)} \boldsymbol{y}_{\mathcal{P}},$$

其中 $L_s^{(2)}$ 和 $L_s^{(3)}$ 分别为 $\frac{1}{2}(\boldsymbol{I}_2+\boldsymbol{J}_2)$ 与 $\frac{1}{3}\boldsymbol{I}_3+\frac{2}{3}\boldsymbol{J}_3$ 的 s 阶乘积. 类似定理 2.4 的证明可证毕.

从图 2.14 可知: 对于较小的 s, $\mathrm{LB}_{\mathrm{LD}}^{(21)}$ 和 $\mathrm{LB}_{\mathrm{LD}}^{(31)}$ 分别要优于 $\mathrm{LB}_{\mathrm{LD}}^{(22)}$ 和 $\mathrm{LB}_{\mathrm{LD}}^{(32)}$.

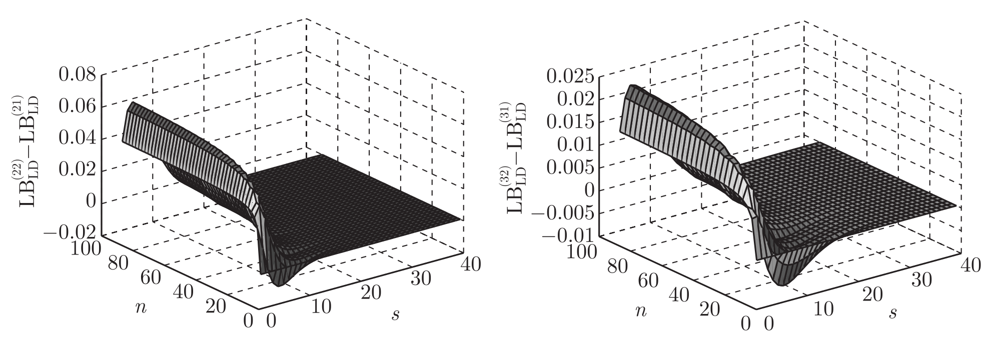


图 2.14　$\mathrm{LB}_{\mathrm{LD}}^{(21)}$ 与 $\mathrm{LB}_{\mathrm{LD}}^{(22)}$, $\mathrm{LB}_{\mathrm{LD}}^{(31)}$ 与 $\mathrm{LB}_{\mathrm{LD}}^{(32)}$ 的比较 (文后附彩图)

对于一般的混水平 U-型设计, Zhou 等 (2008) 对于 $\mathcal{P} \in \mathcal{U}(n; q_1^{s_1} \times \cdots \times q_m^{s_m})$ 获得了 LD 的一些下界, 其中 $q_1, \cdots, q_t$ 为奇数, $q_{t+1}, \cdots, q_m$ 为偶数, $0 \leqslant t \leqslant s$.

定理 2.30　设$\mathcal{P} \in \mathcal{U}(n; q_1^{s_1} \times \cdots \times q_m^{s_m})$, 则

$$\mathrm{LD}^2(\mathcal{P}) \geqslant \mathrm{LB}_{\mathrm{LD}}^{(3)},$$

其中

$$\begin{aligned}\mathrm{LB}_{\mathrm{LD}}^{(3)} =& \Delta_5 + \frac{n-1}{n}\left(\frac{1}{2}\right)^{\sum_{r=t+1}^{m} \frac{ns_r}{q_r(n-1)}} \prod_{r=1}^{t}\left[\prod_{j=1}^{(q_r-1)/2} \frac{q_r+2j-1}{2q_r}\right]^{\frac{2s_rn}{q_r(n-1)}} \\ & \times \prod_{r=t+1}^{m}\left[\prod_{j=1}^{q_r/2-1} \frac{q_r+2j}{2q_r}\right]^{\frac{2s_rn}{q_r(n-1)}},\end{aligned}$$

$$\Delta_5 = \frac{1}{n} - \left(\frac{4}{3}\right)^{\sum_{j=t+1}^{m} s_j} \prod_{j=1}^{t}\left(\frac{3}{4} + \frac{1}{4q_j^2}\right)^{s_j}.$$

定理 2.30 的证明类似于定理 2.11.

当 $q_1 = \cdots = q_m = q$ 时, 定理 2.30 中的下界等价于定理 2.27 中的下界. 当 $q_1 = \cdots = q_t = 3$ 且 $q_{t+1} = \cdots = q_m = 2$ 时, 有如下的定理.

定理 2.31　设$\mathcal{P} \in \mathcal{U}(n; 2^{s_1} \times 3^{s_2})$, 则

$$\mathrm{LD}^2(\mathcal{P}) \geqslant \frac{1}{n} - \left(\frac{3}{4}\right)^{s_1}\left(\frac{7}{9}\right)^{s_2} + \frac{n-1}{n}\left(\frac{1}{2}\right)^{\frac{ns_1}{2(n-1)}}\left(\frac{2}{3}\right)^{\frac{2ns_2}{3(n-1)}}.$$

对于二、三混水平设计, Chatterjee 等 (2012a) 也获得了如下结论, 其下界要优于定理 2.31 中的下界.

定理 2.32　设 $\mathcal{P} \in \mathcal{D}(n; 2^{s_1} \times 3^{s_2})$, 则

$$\mathrm{LD}^2(\mathcal{P}) \geqslant -\left(\frac{3}{4}\right)^{s_1}\left(\frac{7}{9}\right)^{s_2} + \frac{1}{n^2}\left(\frac{1}{2}\right)^{s_1}\left(\frac{2}{3}\right)^{s_2} \sum_{i=0}^{s_1}\sum_{j=0}^{s_2}\binom{s_1}{i}\binom{s_2}{j}\left(\frac{1}{2}\right)^{j}\theta_{ij},$$

其中 θ_{ij} 如定理 2.12 所示.

定理 2.32 的证明类似于定理 2.3.

习　　题

2.1　对于一个单因子试验 $(s=1)$, 证明 n 次试验设计具有最小 $\mathcal{L}_\infty$-星偏差为下面的等间距点集:

$$\mathcal{P} = \left\{\frac{1}{2n}, \frac{3}{2n}, \cdots, \frac{2n-1}{2n}\right\}.$$

2.2　证明具有最小星偏差的一次试验的设计形如 $\mathcal{P}=\{(z,\cdots,z)\}$, z 满足 $z^s+z-1=0$.

2.3　设 $\mathcal{P}=\{\boldsymbol{x}_k=(x_{k1},\cdots,x_{ks}),k=1,\cdots,n\}$ 是单位立方体 $C^s=[0,1]^s$ 上的 n 个点构成的集合. $\mathcal{P}$ 的可卷 L_2-偏差可表示为

$$\mathrm{WD}^2(\mathcal{P})=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\left(\frac{3}{2}\right)^s\sum_{k=1}^n\sum_{l=1}^n\left(\frac{5}{6}\right)^{d_H(k,l)},$$

其中 $d_H(k,l)$ 为 $\boldsymbol{x}_k$ 与 $\boldsymbol{x}_l$ 的 Hamming 距离.

证明：若 D_1 与 D_2 是等价的, 则 $\mathrm{WD}(D_1)=\mathrm{WD}(D_2)$. 两个 U-型设计等价即一个设计可以由另一设计通过交换行或交换列获得.

2.4　在 WD, CD 和 MD 下, 比较下面两个设计的均匀性.

$$\boldsymbol{X}_1=\begin{bmatrix}3/4&3/4&3/4\\3/4&3/4&3/4\\3/4&1/4&3/4\\3/4&3/4&1/4\\1/4&3/4&1/4\end{bmatrix},\quad \boldsymbol{X}_2=\begin{bmatrix}1/4&1/4&1/4\\1/4&3/4&3/4\\3/4&1/4&1/4\\3/4&3/4&1/4\\1/4&3/4&1/4\end{bmatrix}.$$

2.5　编写计算 $\mathrm{WD}(\mathcal{P}),\mathrm{CD}(\mathcal{P}),\mathrm{MD}(\mathcal{P})$ 的 MATLAB 代码, 其中 $\mathcal{P}$ 是 $[0,1]^s$ 上有 n 次试验, s 个因子的设计. 应用你的代码到例 2.1 中的两个设计.

2.6　证明 2.3 节中定义的 CD, WD 及 MD 可表示为 Hamming 距离的函数.

2.7　考虑如下四个设计.

$\mathcal{P}_{5-1}$		$\mathcal{P}_{5-2}$		$\mathcal{P}_{5-3}$		$\mathcal{P}_{5-4}$	
1	2	1	2	1	5	1	1
2	4	2	5	2	1	2	5
3	1	3	3	3	4	3	4
4	3	4	1	4	2	4	3
5	5	5	4	5	3	5	2

计算每个设计的星偏差, CD, WD, MD 及 LD 并展示你的结论.

2.8　算术平均与几何平均间的不等式在推导下界中非常有用. 证明下面的不等式:

设 $a_1,\cdots,a_m$ 为 m 个非负数, 则

$$\bar{a}\equiv\frac{1}{m}\sum_{i=1}^m a_i\geqslant\left[\prod_{i=1}^m a_j\right]^{1/m}\equiv\bar{a}_g,$$

其中 $\bar{a}$ 为 $a_1,\cdots,a_m$ 的算术平均, $\bar{a}_g$ 为 $a_1,\cdots,a_m$ 的几何平均. 上述等式成立当且仅当所有的 a_i 全相等.

2.9　为了便于理解所谓的维数祸根, 考虑 R^n 中球的体积. 一个球心在 R^n 的原点、半径为 r 的球可表示为

$$B_n(r)=\{\boldsymbol{x}:\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=x_1^2+\cdots+x_n^2\leqslant r^2\}.$$

已知球 $B_3(r)$ 的体积为 $\frac{4}{3}\pi r^3$. 一般地

$$\mathrm{Vol}(B_n(r)) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} r^n.$$

显然, 你可能认为 $B_n(r)$ 的体积会变得越来越大. 对于单位球, 它的体积在前面的五个维度中是增加的, 但当 n 趋于无穷时, 体积是下降的. 请研究体积 $B_n(r)$ 随 n 增加时的变化行为.

2.10　在线性内积空间中, 什么是 Cauchy-Schwarz 不等式? 给出证明. 查阅文献找到 Cauchy-Schwarz 不等式更多的应用.

2.11　证明定理 2.30 和定理 2.32.

第 3 章　构造均匀设计的确定性方法

本章和第 4 章将介绍构造均匀设计表的不同方法. 有三种典型的构造均匀设计表的方法:

1) 伪蒙特卡罗法;

2) 组合方法;

3) 数值搜索方法.

本章介绍前两种方法, 而第 4 章介绍第三种方法. 下面网站提供了很多已有的均匀设计表: `http://www.math.hkbu.edu.hk/UniformDesign/`.

3.1 节给出构造均匀设计诸多概念与注意事项, 并给出构造一维和多维均匀设计的简要介绍，并且证明一维均匀设计有一个显式结构. 3.2 节说明对于多因子试验构造均匀设计的复杂度很大. 伪蒙特卡罗方法是构造多因子均匀设计的有效方法. 3.3 节和 3.4 节介绍好格子点法及其推广方法, 3.5 节给出线性水平置换法, 3.6 节给出构造均匀设计的组合方法.

3.1　均匀设计表

为了方便实际应用, 需要不同大小的均匀设计表. 我们有两种途径获得想要的均匀设计: ① 从数据库中查询文献中已获得的许多均匀设计；② 直接由计算机软件搜索得到. 本节中, 我们介绍构造均匀设计表的一些重要事项. 对于单因子试验, 也给出构造均匀设计的理论结果.

3.1.1　均匀设计表的背景

为了构造均匀设计, 需要事先确定一些重要事项, 例如, 试验参数、试验区域、均匀性度量和备选设计集.

试验参数　包括试验次数 n 和因子个数 s.

试验区域　存在不同类型的试验区域 $\mathcal{X}$:

(a) 超立方体: $\mathcal{X}=[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_s,b_s]$ (参笛卡儿积的定义 (1.7)), 因子之间没有约束. 不失一般性, 可以用一个线性映射把 $\mathcal{X}$ 变换到单位超立方体 $C^s=[0,1]^s$.

(b) 单纯形: 在 混料试验 中, 试验区域 $\mathcal{X}$ 为

$$T^s=\{(x_1,\cdots,x_s):x_j\geqslant 0,\ j=1,\cdots,s,\ x_1+\cdots+x_s=1\},\tag{3.1}$$

或

$$T^s(\boldsymbol{a},\boldsymbol{b})=\left\{(x_1,\cdots,x_s):0<a_i\leqslant x_i\leqslant b_i<1,\ i=1,\cdots,s,\ \sum_{i=1}^{s}x_i=1\right\},$$

其中 $\boldsymbol{a}=(a_1,\cdots,a_s)$, $\boldsymbol{b}=(b_1,\cdots,b_s)$.

(c) 格子点集: 若因子为定性因子且 s 个因子的水平分别为 $q_1,\cdots,q_s$, 则试验区域变为

$$\mathcal{X}=\{1,\cdots,q_1\}\times\cdots\times\{1,\cdots,q_s\}, \tag{3.2a}$$

或每个因子都是 q 水平时,

$$\mathcal{X}=\{1,\cdots,q\}^s. \tag{3.2b}$$

注意这里 s 个因子的试验区域都是每个因子的笛卡儿积.

均匀性度量　当试验区域为 C^s 时, 第 2 章已经给出一些均匀性度量. 我们将考虑 2.3.2 小节的中心化偏差, 2.3.3 小节的可卷偏差和 2.3.5 小节的混合偏差, 以及 2.5 节中针对离散点集的离散偏差和 Lee 偏差. 当试验区域为单纯形时, 并不容易选择一个合适的均匀性度量. 第 8 章将讨论这个问题.

备选设计集　设 $\mathcal{D}(n;C^s)$ 为试验区域 C^s 上包含 n 个点的点集的集合. 称 $\mathcal{D}(n;C^s)$ 为备选设计集. 给定某个偏差 D, C^s 上包含 n 个点的均匀设计定义为 $\mathcal{D}(n;C^s)$ 中具有最小偏差值的设计. 自然地, 备选集中的设计个数是无穷的, 因此给定均匀性度量 D, 并不容易找到一个最佳的设计.

为了减少备选设计的个数, 可以把设计限制在格子点上. 假设每个因子有 n 个水平, 则所有的格子点数有 n^s. 记 $D(n;n^s)$ 为从 n^s 个格子点中选取 n 个点组成的设计. 记 $\mathcal{D}(n;n^s)$ 为所有 $D(n;n^s)$ 的集合. 若重复点被允许和不允许时, 则 $\mathcal{D}(n;n^s)$ 的备选设计的个数分别为 n^{sn} 和 $\binom{n^s}{n}$.

显然, $\mathcal{D}(n;n^s)$ 中的备选设计数量仍太大, 以至于即使对于不太大的 n 和 s 时, 计算也不能承受. 因此, 需要进一步减少备选设计数. 文献中, 许多作者已经证明 U-型设计是对均匀设计的一个合理要求. U-型设计见定义 1.6. 当 $n=s$ 时, 一个 U-型设计 $U(n;n^s)$ 为一个 $n\times s$ 的矩阵 $\boldsymbol{X}$, 其每一列都是这 n 个水平 $\left\{\dfrac{2i-1}{2n},i=1,\cdots,n\right\}$的一个置换. 记所有 U-型设计 $U(n;n^s)$ 的集合为 $\mathcal{U}(n;n^s)$, 其中备选设计的个数为 $(n!)^s$.

在实际应用中, 因子的水平数通常小于 n. 对 s 个因子, 设其水平数分别为 $q_1,\cdots,q_s$, 则所有的格子点数为 $N=q_1\cdots q_s$. 由定义 1.6, 备选设计为 $U(n;q_1\times$

$\cdots \times q_s)$. 若 $q_1 = \cdots = q_s = q$, 相应的设计被称为 **对称设计**, 否则为**非对称设计**. 记所有 $U(n; q_1 \times \cdots \times q_s)$ 的集合为 $\mathcal{U}(n; q_1 \times \cdots \times q_s)$, 其中备选设计的个数为 $\prod_{i=1}^s \left[\binom{n}{q_i}\binom{n-q_i}{q_i}\cdots\binom{2q_i}{q_i}\right]$, 这比集合 $\mathcal{U}(n; n^s)$ 中的备选设计的个数小得多. 总之, 备选设计集可以如下依次变小.

$$\boxed{\mathcal{D}(n; C^s)} \Longrightarrow \boxed{\mathcal{D}(n; n^s)} \Longrightarrow \boxed{\mathcal{U}(n; n^s)} \Longrightarrow \boxed{\mathcal{U}(n; q_1 \times \cdots \times q_s)} \tag{3.3}$$

若有些 q_i 相等, 记相应的非对称 U-型设计为 $U(n; q_1^{s_1} \times \cdots \times q_m^{s_m})$, 其中 $s = \sum_{i=1}^m s_i$. 记所有 $U(n; q_1^{s_1} \times \cdots \times q_m^{s_m})$ 和 $U(n; q^s)$ 的集合分别为 $\mathcal{U}(n; q_1^{s_1} \times \cdots \times q_m^{s_m})$ 和 $\mathcal{U}(n; q^s)$. 为简单起见, 有时直接用 $\mathcal{U}$ 表示试验区域.

设计矩阵 一个设计 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\} \subset [0,1]^s$ 可以用一个 $n \times s$ 的矩阵 $\boldsymbol{X}$ 表示, 该矩阵的 n 行 $\boldsymbol{x}_1^{\mathrm{T}}, \cdots, \boldsymbol{x}_n^{\mathrm{T}}$ 分别表示设计的 n 个试验点. 为了强调其为矩阵的形式, 称之为设计矩阵.

根据搜索均匀设计的备选集 (3.3) 的讨论, 可以限制对称设计和非对称设计的设计空间分别为 $\mathcal{U}(n; q^s)$ 或 $\mathcal{U}(n; q_1 \times \cdots \times q_s)$. 在实际应用中, 试验参数可以包括如下几个方面.

试验参数 包括试验次数 n、因子个数 s 和各因子的水平数 $q_1, \cdots, q_s$.

若试验区域为超矩形 $\mathcal{X} = [a_1, b_1] \times \cdots \times [a_s, b_s]$, 由线性变换可把试验区域 $\mathcal{X}$ 变换至 C^s. 考虑构造均匀设计 $U_n(q^s)$, 把 $\{1, \cdots, q\}$ 映射到 $[0,1]$ 有以下几种映射方式:

$$\text{中心化:} \quad x = (u - 1/2)/q, \qquad u = 1, \cdots, q, \tag{3.4a}$$

$$\text{左侧:} \quad x = (u - 1)/q, \qquad u = 1, \cdots, q, \tag{3.4b}$$

$$\text{有端点:} \quad x = (u - 1)/(q - 1), \qquad u = 1, \cdots, q, \tag{3.4c}$$

$$\text{无端点:} \quad x = u/(q + 1), \qquad u = 1, \cdots, q. \tag{3.4d}$$

记 $\boldsymbol{U} = (1, 2, 3, 4)^{\mathrm{T}}$ 为一个四水平、四个试验次数的单因子试验. 图 3.1 给出这四种变换到 $[0,1]$ 的映射方式. 采用中心化映射的设计为$\boldsymbol{X}_{\mathrm{ctr}} = (1/8, 3/8, 5/8, 7/8)^{\mathrm{T}}$, 其具有均匀散布的性质, 每个点分别为区间 $[(i-1)/4, i/4]$ 的中点, $i = 1, \cdots, 4$. 采用左侧映射的设计为 $\boldsymbol{X}_{\mathrm{lft}} = (0, 1/4, 1/2, 3/4)^{\mathrm{T}}$, 其由 $\boldsymbol{X}_{\mathrm{ctr}}$ 平移到左端点处. 采用有端点的映射方式时, $\boldsymbol{X}_{\mathrm{ext}} = (0, 1/3, 2/3, 1)^{\mathrm{T}}$, 其间隔也均匀且有两个端点. 采用无端点的映射方式的设计为 $\boldsymbol{X}_{\mathrm{mis}} = (1/5, 2/5, 3/5, 4/5)^{\mathrm{T}}$, 其和 $\boldsymbol{X}_{\mathrm{ext}}$ 类似, 但是无端点. 在实际应用中, 中心化映射 $x = (u - 1/2)/q (u = 1, \cdots, q)$ 最常用.

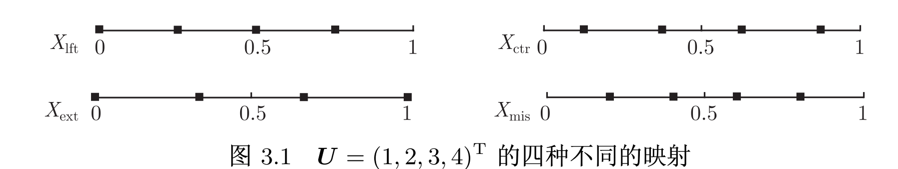


图 3.1　$\boldsymbol{U} = (1, 2, 3, 4)^{\rm T}$ 的四种不同的映射

例 3.1　为了解释 (3.4) 中的四种映射的特点, 考虑钟摆周期和振幅之间的函数关系这一简单的试验. 我们希望在钟摆处于不同振幅时, 测量钟摆周期, 并对响应 (周期) 和因子 (振幅) 之间建立模型. 试验区域为 $[0°, 180°]$, 当处于 0° 时, 钟摆的开始位置处于垂直状态, 而处于 180° 时, 钟摆的开始位置处于水平状态. 对于设计 $\boldsymbol{U} = (1, 2, 3, 4)^{\rm T}$, 根据映射 (3.4), 对设计变换到 $[0°, 180°]$ 内存在下面四种映射方式:

$$\boldsymbol{X}_{\rm lft} = (0°,\ 45°,\ 90°,\ 135°)^{\rm T}, \tag{3.5a}$$

$$\boldsymbol{X}_{\rm ctr} = (22.5°,\ 67.5°,\ 112.5°,\ 157.5°)^{\rm T}, \tag{3.5b}$$

$$\boldsymbol{X}_{\rm ext} = (0°,\ 60°,\ 120°,\ 180°)^{\rm T}, \tag{3.5c}$$

$$\boldsymbol{X}_{\rm mis} = (36°,\ 72°,\ 108°,\ 144°)^{\rm T}. \tag{3.5d}$$

3.1.2 小节将说明中心化映射 (3.4a) 是最常用的选择. 由该映射, n 行、s 列的设计矩阵 $\boldsymbol{U}$ 被映射到 $[0,1]^s$ 中的点集 $\boldsymbol{X}$ (为了强调 $\boldsymbol{U}$, 也可以记为 $\boldsymbol{X}_{\boldsymbol{U}}$). 类似的变换也可应用于非对称设计.

定义 3.1　对任意设计 $\boldsymbol{U} \in \mathcal{U}(n; q_1 \times \cdots \times q_s)$, 对第 i 列考虑映射 $f: u \to (2u-1)/(2q_i), u = 1, \cdots, q_i$, $i = 1, \cdots, s$, 则得到的设计矩阵 $\boldsymbol{X}$ 被称为 $\boldsymbol{U}$ 的导出矩阵. 给定均匀性度量 D, 定义 $D(\boldsymbol{U}) = D(\boldsymbol{X})$.

定义 3.1 中的映射是一一对应的.

例 3.2　给出下面的 U-型对称设计 $U(12; 12^4)$

$$\boldsymbol{U} = \begin{bmatrix} 1 & 10 & 4 & 7 \\ 2 & 5 & 11 & 3 \\ 3 & 1 & 7 & 9 \\ 4 & 6 & 1 & 5 \\ 5 & 11 & 10 & 11 \\ 6 & 9 & 8 & 1 \\ 7 & 4 & 5 & 12 \\ 8 & 2 & 3 & 2 \\ 9 & 7 & 12 & 8 \\ 10 & 12 & 6 & 4 \\ 11 & 8 & 2 & 10 \\ 12 & 3 & 9 & 6 \end{bmatrix},$$

相应的导出矩阵为

$$\boldsymbol{X}=\begin{pmatrix} 1/24 & 19/24 & 7/24 & 13/24 \\ 3/24 & 9/24 & 21/24 & 5/24 \\ 5/24 & 1/24 & 13/24 & 17/24 \\ 7/24 & 11/24 & 1/24 & 9/24 \\ 9/24 & 21/24 & 19/24 & 21/24 \\ 11/24 & 17/24 & 15/24 & 1/24 \\ 13/24 & 7/24 & 9/24 & 23/24 \\ 15/24 & 3/24 & 5/24 & 3/24 \\ 17/24 & 13/24 & 23/24 & 15/24 \\ 19/24 & 23/24 & 11/24 & 7/24 \\ 21/24 & 15/24 & 3/24 & 19/24 \\ 23/24 & 5/24 & 17/24 & 11/24 \end{pmatrix}=\begin{pmatrix} 0.042 & 0.792 & 0.292 & 0.542 \\ 0.125 & 0.375 & 0.875 & 0.208 \\ 0.208 & 0.042 & 0.542 & 0.708 \\ 0.292 & 0.458 & 0.042 & 0.375 \\ 0.375 & 0.875 & 0.792 & 0.875 \\ 0.458 & 0.708 & 0.625 & 0.042 \\ 0.542 & 0.292 & 0.375 & 0.958 \\ 0.625 & 0.125 & 0.208 & 0.125 \\ 0.708 & 0.542 & 0.958 & 0.625 \\ 0.792 & 0.958 & 0.458 & 0.292 \\ 0.875 & 0.625 & 0.125 & 0.792 \\ 0.958 & 0.208 & 0.708 & 0.458 \end{pmatrix}.$$

该导出矩阵的平方偏差值为 $\mathrm{CD}^2(\boldsymbol{X})=0.0114, \mathrm{WD}^2(\boldsymbol{X})=0.0339$ 且 $\mathrm{MD}^2(\boldsymbol{X})=0.0386$. 根据定义 3.1, 设计 $\boldsymbol{U}$ 的偏差等于导出矩阵 $\boldsymbol{X}$ 的偏差.

定义 3.2 给定备选设计空间 $\mathcal{U}$, 若设计 $\mathcal{P}^*$ 在 $\mathcal{U}$ 上有最小的偏差度量 D 值, 则称其为均匀设计, 即

$$D(\mathcal{P}^*)=\min_{\mathcal{P}\in\mathcal{U}} D(\mathcal{P}). \tag{3.6}$$

一般地, 设计空间 $\mathcal{U}$ 可选为 $\mathcal{D}(n;C^s)$, U-型设计 $\mathcal{U}(n;q^s)$ 或 $\mathcal{U}(n;q_1\times\cdots\times q_s)$. $\mathcal{U}(n;q^s)$ 和 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 中的均匀设计分别记为 $U_n(q^s)$ 和 $U_n(q_1\times\cdots\times q_s)$. 类似于正交设计表, 均匀设计也可以由表格的形式表示, 称之为*均匀设计表*.

根据 2.2.2 小节中对均匀性测度的要求准则 C_1, 对设计表做行置换和列置换都不影响其偏差值, 因此当 $s>1$ 时, 最小化问题 (3.6) 往往有很多解. 同时, 许多偏差也具有对于设计区域中心点的反射不变性. 称具有相同偏差值的设计为*等价设计*. 一个搜索均匀设计的有效算法在搜索过程中应该避免考虑与给定设计等价的其他设计.

一个设计若近似地解决最小化问题 (3.6), 即其偏差值已经非常接近于最小值, 则称该设计为*低偏差设计*或*近似均匀设计*. 需要注意的是, 文献中很多设计都是近似均匀设计. 由于优化问题 (3.6) 往往很难找到其最优解, 近似均匀设计也是可以接受的, 其也常被认为是均匀设计.

3.1.2 单因子均匀设计

具有连续试验区域的单因子试验, 不失一般性, 考虑其试验区域为 $[0,1]$. 本小

节将说明在中心化偏差、可卷偏差、混合偏差和星偏差意义下, 单因子均匀设计可以用解析的方式给出.

定理 3.1 考虑单因子设计的备选集为 $\mathcal{D}(n;[0,1])$. 在不同均匀性准则下的单因子均匀设计分别如下所示.

(a) (Fang and Wang, 1994) 在星偏差意义下, $[0,1]$ 上的单因子均匀设计为

$$\boldsymbol{X}^* = \left(\frac{1}{2n}, \frac{3}{2n}, \cdots, \frac{2n-1}{2n}\right)^{\mathrm{T}}, \tag{3.7}$$

其星偏差值为 $D^*(X^*) = 1/(2n)$.

(b) (Fang et al., 2002a) 设计 (3.7) 也是在中心化偏差意义下唯一的均匀设计, 其中心化偏差值的平方为 $\mathrm{CD}^2(\boldsymbol{X}^*) = 1/(12n^2)$.

(c) (Fang and Ma, 2001b) 在可卷偏差意义下, [0,1] 上的单因子均匀设计为

$$\boldsymbol{X}_\delta = \left(\frac{\delta}{n}, \frac{1+\delta}{n}, \cdots, \frac{n-1+\delta}{n}\right)^{\mathrm{T}}, \tag{3.8}$$

其中任意实数 $\delta \in [0,1]$, 且其可卷偏差值的平方为$\mathrm{WD}^2(\boldsymbol{X}_\delta) = 1/(6n^2)$.

(d) (Zhou et al., 2013) 设计 (3.7) 也是在混合偏差意义下唯一的均匀设计, 其混合偏差值的平方为 $\mathrm{MD}^2(\boldsymbol{X}^*) = 1/(8n^2)$.

证明 我们只证明定理 3.1(d). 定理 3.1(b) 和 (c) 的证明类似, 而定理 3.1(a) 的证明可参见 (Fang and Wang, 1994) 的第 16-17 页. 设 $\mathcal{P} = \{x_1, x_2, \cdots, x_n\}$ 为 [0,1] 上的集合. 不失一般性, 设 $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n$ 且记 $y_k = x_k - 1/2$, $k = 1, 2, \cdots, n$. 则

$$\begin{aligned}
&\mathrm{MD}^2(\mathcal{P})\\
=&\frac{19}{12} - \frac{2}{n}\sum_{k=1}^{n}\left(\frac{5}{3} - \frac{|y_k|}{4} - \frac{|y_k|^2}{4}\right)\\
&+ \frac{1}{n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\left(\frac{15}{8} - \frac{|y_k|}{4} - \frac{|y_j|}{4} - \frac{3|y_k - y_j|}{4} - \frac{|y_k - y_j|^2}{2}\right)\\
=&\frac{19}{12} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\left[\left(\frac{15}{8} - \frac{|y_k|}{4} - \frac{|y_j|}{4} - \frac{3|y_k - y_j|}{4} - \frac{|y_k - y_j|^2}{2}\right)\right.\\
&\left. - \left(\frac{5}{3} - \frac{|y_k|}{4} - \frac{|y_k|^2}{4}\right) - \left(\frac{5}{3} - \frac{|y_j|}{4} - \frac{|y_j|^2}{4}\right)\right]\\
=&\frac{1}{8} + \frac{1}{2n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\left((y_i - y_k)^2 - |y_i - y_k|\right) + \frac{1}{2n}\left(\sum_{k=1}^{n} y_k^2 - \frac{1}{n}\sum_{k>j}(y_k - y_j)\right)
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{8}+\frac{1}{2n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\left(|y_i-y_k|-\frac{|k-j|}{n}\right)^2+\frac{1}{2}\left(\frac{1}{6n^2}-\frac{1}{6}\right)\\
&+\frac{1}{2n}\sum_{k=1}^{n}\left(y_k-\frac{2k-1-n}{2n}\right)^2+\frac{1}{2}\left(\frac{1}{12n^2}-\frac{1}{12}\right)\\
=&\frac{1}{8n^2}+\frac{1}{2n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\left(|y_i-y_k|-\frac{|k-j|}{n}\right)^2+\frac{1}{n}\sum_{k=1}^{n}\left(y_k-\frac{2k-1-n}{2n}\right)^2\\
\geqslant&\frac{1}{8n^2}.
\end{aligned}
$$

因此, $\mathrm{MD}^2(\mathcal{P})$ 取到最小值当且仅当$y_k=\dfrac{2k-1-n}{2n}$, 即 $x_k=\dfrac{2k-1}{2n}, k=1,2,\cdots,n$. 定理得证.

定理 3.1(a), (b) 和 (d) 说明了中心化映射 (3.4a) 的合理性, 即在该映射下, n 个试验次数的单因子 U-型设计 $\boldsymbol{U}=(1,\cdots,n)^{\mathrm{T}}$ 可以最小化星偏差、中心化偏差和混合偏差. 然而, 定理 3.1(c) 说明把设计点整体平移, 其可卷偏差并不改变. 因此, 其均匀设计并不唯一. 显然, 对应于中心化映射, 设计矩阵 $\boldsymbol{X}_\delta$ 中的 $\delta=1/2$. 定理 3.1 也说明在不同的偏差下, 其相应的均匀设计可能不一样.

在实际中, 常考虑中心化映射, 因为单因子设计 $\boldsymbol{X}_{\mathrm{ctr}}$ 有最小的中心化偏差、混合偏差, 也有最小的可卷偏差. 中心化映射带来的不便之处在于其真实水平可能不是一个整数, 这会给试验人员带来不便. 例如, 对于温度而言, $\boldsymbol{X}_{\mathrm{ctr}}$ 的对应真实水平数可能是分数, 而其他映射下的真实水平为整数. 设计 $\boldsymbol{X}_{\mathrm{lft}}$ 有最小的可卷偏差值, 但其中心化偏差和混合偏差并不是最小的.

设计 $\boldsymbol{X}_{\mathrm{ext}}$ 包括因子的两个端点, 其优点是只需要对拟合模型做内插估计, 而不需要做外插估计. 同时, 其真实水平可能是整数. 然而, $\boldsymbol{X}_{\mathrm{ext}}$ 的一个缺点是, 在有些应用中, 试验不能取到边界值. 例如, 在例 3.1 的钟摆试验中, 当振幅为 0° 时, 试验是无法进行的. 类似于 $\boldsymbol{X}_{\mathrm{ext}}$, 设计 $\boldsymbol{X}_{\mathrm{mis}}$ 的真实水平可能是整数. $\boldsymbol{X}_{\mathrm{mis}}$ 舍弃试验区域的两个端点, 其可以避免在边界处无法做试验的情形.

设计 $\boldsymbol{X}_{\mathrm{lft}}$ 和 $\boldsymbol{X}_{\mathrm{ctr}}$ 是两类*等距设计*, 且两者之间可以通过平移得到, 因此也称之为*平移设计*, 见 (3.8) 式. 当 $\delta=0$ 时, 设计为 $\boldsymbol{X}_{\mathrm{lft}}$; 当 $\delta=1/2$ 时, 设计为 $\boldsymbol{X}_{\mathrm{ctr}}$.

选取随机数 δ 可以消除设计的偏差. 例如, 由随机设计点抽样而得到的某函数的样本均值的期望值, 恰好等于该函数的总体均值. 该结论对于确定性的设计, 不一定成立. 然而, 对于任意平移 δ, 其真实水平可能不是整数. 这个性质对于实验室试验可能会带来不便, 但是对于计算机试验而言并无影响.

3.2　多因子均匀设计

定理 3.1 给出单因子均匀设计的结论, 然而该结果并不能推广至多因子情形. 本节将讨论多因子情形下构造方法的复杂度和介绍一些构造方法.

3.2.1　构造方法的复杂度

在 3.1.1 小节中, 我们指出构造多因子均匀设计的复杂度. 首先典型设计空间 $\mathcal{D}(n;[0,1]^s)$ 太过复杂，且在 $\mathcal{D}(n;[0,1]^s)$ 上的偏差函数往往是多峰的, 这导致传统的优化算法并不适用. 为了避免困难, 选取具有良好结构的子集作为候选设计集. 由 (3.3), 建议考虑用 U-型设计集 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 或 $\mathcal{U}(n;q^s)$ 作为候选集.

定义 3.2 说明寻找一个均匀设计实际上是优化问题 (3.6), 其中 $\mathcal{U}$ 为候选设计集. 为了求解优化问题, 下面考虑两种方法.

(a) *理论方法*. 对于单因子试验, 3.1 节已经给出基于理论方法的不同偏差下的均匀设计. 对于多因子试验, 也有一些生成均匀设计的理论方法. 在 3.3 节介绍由数论专家 Korobov (1959) 提出的好格子点法 (good lattice point method, GLPM). Korobov 给出了该方法的理论性质并提供一批表格以供实际应用. 好格子点法是一种用于构造低偏差点集的方法, 在伪蒙特卡罗方法和均匀设计的构造方法起到重要作用. 另一种构造均匀设计的理论方法是根据组合设计和均匀设计之间的关系而得到的. 3.6 节将介绍组合构造方法并应用组合设计方法来构造均匀设计.

(b) *数值优化算法*. 目前为止, 大部分的均匀设计都是由数值优化算法得到的. 传统的优化算法并不适合用于搜索均匀设计, 因为偏差函数往往是多峰的. 若考虑设计空间为 $\mathcal{D}(n;[0,1]^s)$, 搜索均匀设计相当于在 ns 维实空间 R^{ns} 上的搜索一个最优点. 在 (3.6) 中, 目标函数 D 有许多局部最小值. 因此, 即使当 n 和 s 不太大时, 也很难用优化算法搜索到最小偏差值的全局最优设计. 例如, 当 $n>10$ 时, 已经很难用优化算法搜索到最优的设计了. 因此, 好格子点法和优化算法都建议减少设计空间. 正如 (3.3) 中建议, U-型设计是合适的备选设计空间.

例 3.3　*考虑当 $2\leqslant n\leqslant 13$ 时, 分别比较在备选设计空间 $\mathcal{D}(n;[0,1]^2)$ 和 $\mathcal{U}(n;n^2)$ 上的均匀设计. 表 3.1 和表 3.2 给出在中心化偏差下, $\mathcal{D}(n;[0,1]^2)$ 和 U-型设计集 $\mathcal{U}(n;n^2)$ 上的二因子均匀设计表, 分别记为 $\boldsymbol{U}^*$ 和 $\boldsymbol{U}_0^*$. 这些均匀 U-型设计 $\boldsymbol{U}^*$ 如图 3.2 所示. 为表示方便, 给出这些 U-型设计表, 其偏差值由定义 3.1 通过导出矩阵 $\boldsymbol{X}^*$ 得到. 类似地, $\boldsymbol{U}_0^*$ 的偏差值也是由其导出矩阵 $\boldsymbol{X}_0^*=(\boldsymbol{U}_0^*-0.5)/n$ 而得到的. 易知这两类均匀设计和其中心化偏差值都非常接近. 因此, 构造均匀设计时, 可以合理地限制其搜索空间为 $\mathcal{U}(n;n^s)$. 此外, U-型设计还有一个优点是可以均匀地设定水平值, 给试验人员带来便利.*

表 3.1 比较不同备选设计集的二因子均匀设计

试验次数	$\boldsymbol{U}^* \in \mathcal{U}(n;n^2)$			$\boldsymbol{U}_0^* \in \mathcal{D}(n;[0,1]^2)$		
n	u_{i1}	u_{i2}	偏差	u_{i1}	u_{i2}	偏差
2	1	2	6.228×10^{-2}	1.062	1.938	5.989×10^{-2}
	2	1		1.938	1.062	
3	1	2	2.958×10^{-2}	1.000	2.000	2.884×10^{-2}
	2	1		2.000	1.000	
	3	3		2.910	2.911	
4	1	3	1.626×10^{-2}	1.033	2.992	1.619×10^{-2}
	2	1		2.008	1.033	
	3	4		2.992	3.967	
	4	2		3.967	2.008	
5	1	4	1.105×10^{-2}	1.055	3.986	1.094×10^{-2}
	2	1		2.014	1.055	
	3	3		3.000	3.000	
	4	5		3.985	4.945	
	5	2		4.945	2.015	
6	1	4	7.628×10^{-3}	1.021	4.010	7.612×10^{-3}
	2	2		1.985	1.986	
	3	6		2.990	5.978	
	4	1		4.010	1.022	
	5	5		5.015	5.014	
	6	3		5.978	2.989	
7	1	3	5.824×10^{-3}	1.036	3.039	5.782×10^{-3}
	2	6		1.991	6.009	
	3	1		3.039	1.036	
	4	4		4.000	4.000	
	5	7		4.961	6.963	
	6	2		6.009	1.991	
	7	5		6.963	4.961	
8	1	3	4.475×10^{-3}	1.053	2.976	4.456×10^{-3}
	2	7		1.996	7.005	
	3	5		3.019	4.987	
	4	1		3.980	1.017	
	5	8		5.019	7.983	
	6	4		5.980	4.013	
	7	2		7.004	1.996	
	8	6		7.947	6.023	
9	1	4	3.583×10^{-3}	1.029	4.011	3.575×10^{-3}
	2	7		1.983	7.001	
	3	2		2.999	1.983	
	4	9		4.010	8.971	
	5	5		5.000	5.000	
	6	1		5.990	1.028	
	7	8		7.001	8.017	
	8	3		8.017	2.999	
	9	6		8.971	5.990	

表 3.2　比较不同备选设计集的二因子均匀设计

试验次数 n	$\boldsymbol{U}^* \in \mathcal{U}(n; n^2)$ u_{i1}	u_{i2}	偏差	$\boldsymbol{U}_0^* \in \mathcal{D}(n; [0,1]^s)$ u_{i1}	u_{i2}	偏差
10	1	8	2.953×10^{-3}	1.072	8.031	2.937×10^{-3}
	2	2		2.004	2.002	
	3	6		3.014	5.996	
	4	4		4.003	4.002	
	5	10		4.974	9.986	
	6	1		6.026	1.011	
	7	7		6.997	6.999	
	8	5		7.984	5.005	
	9	9		8.996	8.996	
	10	3		9.928	2.970	
11	1	7	2.467×10^{-3}	1.022	6.973	2.453×10^{-3}
	2	2		2.006	2.007	
	3	10		2.983	10.011	
	4	4		4.011	4.010	
	5	9		5.046	8.970	
	6	5		6.000	5.000	
	7	1		7.048	1.024	
	8	11		8.022	10.949	
	9	6		8.999	5.962	
	10	3		10.013	2.990	
	11	8		10.950	8.022	
12	1	6	2.076×10^{-3}	1.018	5.973	2.065×10^{-3}
	2	10		1.994	10.055	
	3	2		2.946	1.993	
	4	8		3.997	8.014	
	5	4		4.987	3.997	
	6	12		5.972	11.988	
	7	1		7.027	1.012	
	8	9		8.014	9.004	
	9	5		9.004	4.987	
	10	11		10.054	11.008	
	11	3		11.007	2.945	
	12	7		11.988	7.028	
13	1	6	1.747×10^{-3}	1.016	6.007	1.744×10^{-3}
	2	11		1.995	11.024	
	3	2		2.975	1.994	
	4	9		4.002	8.979	
	5	4		5.021	4.002	
	6	13		6.007	12.980	
	7	7		7.000	7.000	
	8	1		7.992	1.017	
	9	10		8.979	9.998	
	10	5		9.997	5.020	
	11	12		11.024	12.005	
	12	3		12.005	2.976	
	13	8		12.979	7.992	

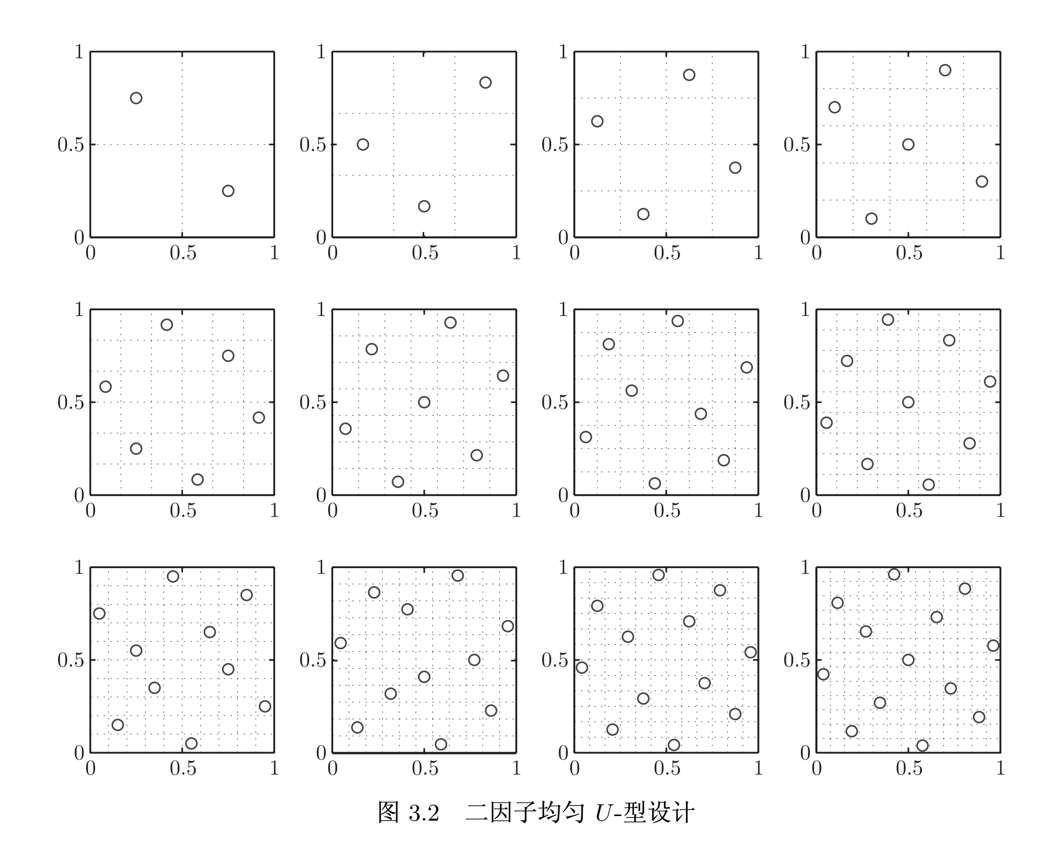


图 3.2 二因子均匀 U-型设计

3.2.2 构造方法注解

从现在开始, 选择 U-型设计集作为设计空间. 易知, U-型设计空间可以看为 R^{ns} 中的离散点集. 因此, 并不存在目标函数的连续结构, 也更不存在其导数了, 因此不能使用传统的优化算法. 为了求解 (3.6), 需要新的优化技术, 例如, 用启发式全局优化算法在候选设计空间中搜索近似均匀设计. 文献中存在诸多启发式全局优化算法, 例如, 门限接受法、模拟退火算法、遗传算法等. 用这些算法得到的设计并不一定能保证具有最小的偏差值, 但可以用较小的代价来获得低偏差设计. 为方便记，也称由这些算法得到的低偏差设计为均匀设计, 即使其可能达不到最小偏差值.

在理论研究和选择数值优化算法中, 需要注意下面的一些事情.

i) 需要考虑合适的候选设计集. U-型设计或其子集是合适的选择.

ii) 设计的行置换或列置换不改变偏差值, 且许多偏差关于试验区域的中心点具有反射不变性. 因此, 在候选设计集中需要能鉴别出这种等价设计.

iii) 需要给出一个好的搜索算法. 一般而言, 没有一个算法在所有情形下都是

最优的. 选择一个好算法依赖于候选集的大小. 对于一个小的候选集, 可以用穷举法搜索. 对于一个更大的候选集, 需要考虑一个随机搜索算法, 例如, 门限接受法或进化算法等. 有时, 可以从一个低偏差的设计的投影或该设计的部分点而得到一个好设计.

iv) 通常一个均匀设计每个因子的水平数和其试验次数一样大, 不过在很多实体试验中, 水平数不易太大. 在某种程度上, 这会增加搜索算法的复杂度.

下面的注释对构造均匀设计可能面临的问题会有所帮助.

注 1 第 2 章给出了不同种类的偏差. 幸运的是, 不同的偏差下相应的均匀设计的性质是类似的, 并没有很大的改变. 换句话说, 一个设计在某个偏差准则下具有较小的偏差值, 则在另一个偏差准则下也会有相对较小的偏差值. 本章中, 常用中心化偏差区别不同设计.

注 2 实验室试验的次数通常不会太大, 比如几十次试验. 不过计算机试验的次数可以达到几千甚至几百万之巨. 实验室试验和计算机试验这两类试验的候选设计集可以类似地构造, 不过搜索均匀设计的方法有较大区别. 对于试验次数较小的情形, 我们可以尝试搜索一个具有最小偏差的设计, 因为每个试验点都很重要. 对于试验次数较大的情形, 通常选一个容易计算的均匀性测度, 且把搜索空间限制在候选设计集的子集上.

注 3 容易证明在 U-型设计集上得到的均匀设计可能不是 $\mathcal{D}(n;[0,1]^s)$ 上的均匀设计. 不过在 U-型设计集上得到的均匀设计的偏差值往往非常接近于$\mathcal{D}(n;[0,1]^s)$ 上的均匀设计的偏差值.

在 U-型设计集 $\mathcal{U}(n;n^s)$ 中候选设计数为 $(n!)^s$. 不失一般性, 可选择设计的第一列为 $(1,\cdots,n)^{\mathrm{T}}$, 因为可以对任意设计做行置换可使得第一列恰好为 $(1,\cdots,n)^{\mathrm{T}}$ 且其偏差值不变. 则候选设计数量可以降低至 $(n!)^{s-1}$. 若 $(n!)^{s-1}$ 不太大, 可以用穷举法寻找最优的 U-型设计. 否则, 需要再次减少设计空间. 文献中, 好格子点法是第一种这类方法, 3.3 节将介绍好格子点法及其推广.

3.3 好格子点法及其推广

好格子点法在伪蒙特卡罗法中是常用的方法. 该方法首先由 Korobov (1959) 提出, 用于多维积分的数值逼近. 更多的讨论见 (Hua and Wang, 1981; Shaw, 1988; Fang and Wang, 1994). 本节中, 我们介绍好格子点法及其推广方法.

3.3.1 好格子点法

设 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 为两个设计. 若 $\boldsymbol{X}_1$ 可由 $\boldsymbol{X}_2$ 的行置换和/或列置换而得到, 则 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 相互等价. 若把等价的 U-型设计记为一个设计, 则在 $\mathcal{U}(n;n^s)$ 中至多

有 $(n!)^{s-1}$ 个不等价的 $U(n;n^s)$ 设计. 然而, 即使对于一般大小的试验次数和因子个数, $(n!)^{s-1}$ 的数值也太大了. 因此, 还需要减少设计空间的复杂度. Fang (1980) 与 Fang 和 Wang (1981) 首先把好格子点法应用到均匀设计的构造中.

好格子点法的目的是生成一个合适的 $\mathcal{U}(n;n^s)$ 的子集. 好格子点法的关键步骤是生成 U-型设计的第一行, 称其为**生成向量**, 并记为 $\boldsymbol{h}=(h_1,\cdots,h_s)$. 则设计的第 j 列为

$$\boldsymbol{h}_j=(h_j,2h_j,\cdots,nh_j)^{\mathrm{T}}\ (\widetilde{\mathrm{mod}}\ n),\quad j=1,2,\cdots,s \tag{3.9}$$

其中算子 $\widetilde{\mathrm{mod}}$ 为 1.2 节中定义的特殊模运算, 该运算常应用于好格子点法中. 显然, 每个 $\boldsymbol{h}$ 要满足以下条件:

(a) h_i 为正整数且 $h_i<n$, $i=1,\cdots,s$;

(b) 每列 $\boldsymbol{h}_j$ 为 $\{1,2,\cdots,n\}$ 的一个置换;

(c) 每个 h_i 都是不同的, 则可假设 $1\leqslant h_1<h_2\leqslant\cdots<h_s<n$;

(d) 矩阵 $(\boldsymbol{h}_1,\cdots,\boldsymbol{h}_s)$ 是列满秩的, 即其秩为 s.

注意 $\boldsymbol{h}_j=(h_j,2h_j,\cdots,nh_j)^{\mathrm{T}}\ (\widetilde{\mathrm{mod}}\ n)$ 可能是 $\{1,2,\cdots,n\}$ 的置换, 也可能不是. 例如, 当 $n=8$, $h_2=2$ 和 $h_3=3$ 时, 可得两个列向量: $(h_2,2h_2,\cdots,8h_2)=(2,4,6,8,2,4,6,8)(\widetilde{\mathrm{mod}}\ 8)$ 且 $(h_3,2h_3,\cdots,8h_3)=(3,6,1,4,7,2,5,8)(\widetilde{\mathrm{mod}}\ 8)$. 第二个列向量是 $\{1,2,\cdots,8\}$ 的置换, 而第一个列向量不是. 保证条件 (b) 成立的充要条件是 n 和 h 的最大公因数为 1, 即 $\gcd(n,h)$=1. 通常, 可以取 $h_1=1$.

设 m 为满足条件 (a)—(c) 的 h_j 的个数. 由数论理论可知, $m=\phi(n)$, 其中 $\phi(\cdot)$ 为**欧拉函数**(Hua and Wang, 1981), 具体定义如下. 对于任意一个正整数 n, 存在一个唯一的素数分解 $n=p_1^{r_1}\cdots p_t^{r_t}$, 其中 $p_1,\cdots,p_t$ 为不同的素数, 且 $r_1,\cdots,r_t$ 为正整数. 则欧拉函数为

$$\phi(n)=n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_t}\right). \tag{3.10}$$

当 n 为素数时, 则 $\phi(n)=n-1$. 例如, $\phi(31)=30$, 因为 31 是素数; 而 $30=2\times3\times5$, 则 $\phi(30)=30(1-1/2)(1-1/3)(1-1/5)=8$. 当 $s>m=\phi(n)$ 时, 不能由好格子点法得到 U-型设计 $U(n;n^s)$; 当 $s<m$ 时, 存在 $\begin{pmatrix} m\\ s\end{pmatrix}$ 个可能不同的 U-型设计 $U(n;n^s)$, 不过其中有些设计的列秩可能小于 s. 记正整数候选集

$$\mathcal{H}_n=\{h:h<n,n\ \text{和}\ h\ \text{互质}\}.$$

若 $h\in\mathcal{H}_n$, 则 $n-h\in\mathcal{H}_n$. 易知, 由 h 和 $n-h$ 分别生成的列 $\boldsymbol{h}_h$ 和 $\boldsymbol{h}_{n-h}$, 有性质 $\boldsymbol{h}_h+\boldsymbol{h}_{n-h}=(n,\cdots,n,2n)'$. 进一步地, 若 $\tilde{h}\in\mathcal{H}_n$ 且 $\tilde{h}$ 不等于 h 或 $n-h$,

则 $\boldsymbol{h}_h + \boldsymbol{h}_{n-h} = \boldsymbol{h}_{\tilde{h}} + \boldsymbol{h}_{n-\tilde{h}}$. 因此, 元素 $\{h, n-h, \tilde{h}, n-\tilde{h}\}$ 不能同时出现在生成向量, 否则生成的设计矩阵不能为列满秩的. 所以, 好格子点法能生成的列数被限制为

$$k(n) = \frac{\phi(n)}{2} + 1, \tag{3.11}$$

其意味着, 生成向量至多有一组 $\{h, n-h\}$.

设 $1 = h_1 < h_2 < \cdots < h_{\phi(n)/2}$ 都来自于 $\mathcal{H}_n$. 在 $\phi(n)/2$ 组 $\{h_i, n-h_i\}, i = 1, \cdots, \phi(n)/2$ 里面, 可以随机选择一组 $\{h_i, n-h_i\}$, 并随机地从其他 $\phi(n)/2-1$ 组中各任选一个元素, 组成 $k(n)$ 个元素的生成向量的候选集, 其可以保证得到的好格子点为列满秩. 总共有 $2^{\phi(n)/2-1}\phi(n)/2$ 种选择, 且对于每种选择可产生 $\begin{pmatrix} k(n) \\ s \end{pmatrix}$ 个 U-型设计. 因此, 在设计空间中总共有 $\phi(n)2^{\phi(n)/2-2}\begin{pmatrix} k(n) \\ s \end{pmatrix}$个 U-型设计, 记为 $\mathcal{G}_{n,s}$. $\mathcal{G}_{n,s}$ 中的候选设计个数远远小于 $\mathcal{U}(n; n^s)$ 中的设计个数. 这是好格子点法的一个优点. 在 $\mathcal{G}_{n,s}$ 中偏差最小的设计是一个近似均匀设计. 好格子点法的具体算法如下所示.

算法 3.1 (好格子点法)

步骤 1 给定参数 n 和 s, 确定正整数候选集 $\mathcal{H}_n$. 当 $s \leqslant k(n)$ 时, 转步骤 2, 否则该方法不能生成近似均匀设计.

步骤 2 设 $\boldsymbol{h} = (h_1, \cdots, h_s)$ 的元素为 $\mathcal{H}_n$ 中 s 个互不相同的数, 则得到 $n \times s$ 矩阵 $\boldsymbol{U} = (u_{ij})$ 其中 $u_{ij} = ih_j \ (\widetilde{\text{mod}}\ n)$. 简记 $\boldsymbol{U}(n, \boldsymbol{h})$ 为 $\boldsymbol{U}$, 其中 $\boldsymbol{h}$ 为 $\boldsymbol{U}$ 的生成向量. 记 $\mathcal{G}_{n,s}$ 为全体具有秩为 s 的 $\boldsymbol{U}(n, \boldsymbol{h})$.

步骤 3 寻找一个生成向量 $\boldsymbol{h}^*$ 使得其相应的 $\boldsymbol{U}(n, \boldsymbol{h}^*)$ 在 $\mathcal{G}_{n,s}$ 中具有最小的偏差. 称 $\boldsymbol{U}(n, \boldsymbol{h}^*)$ 为一个 (近似) 均匀设计 $U_n(n^s)$.

由好格子点法生成的设计也被称为*好格子点*. 实际上, 可以进一步地降低好格子点法的计算复杂度:

(a) 对于任意整数 h, 若存在 h^* 使得 $hh^* = 1 \ (\text{mod}\ n)$, 则称 h^* 为 h 在模 n 意义下的*逆*, 并记为 h^{-1}. 若 h^{-1} 存在, 则易知 $h^{-1} \in \mathcal{H}_n$ 当且仅当 $h \in \mathcal{H}_n$.

(b) 对每个生成向量 $\boldsymbol{h} = (h_1, \cdots, h_s)$, 其中 $h_1 < \cdots < h_s$, 若 $h_1 \neq 1$, 第一列 $\boldsymbol{h}_1 = (h_1, 2h_1, \cdots, nh_1) \ (\widetilde{\text{mod}}\ n)$ 为 $\{1, 2, \cdots, n\}$ 的一个置换. 则可以对设计进行行置换使其第一列为 $(1, 2, \cdots, n)^{\mathrm{T}}$. 则置换后的设计也是另一个好格子点设计, 其生成向量为 $\tilde{\boldsymbol{h}} = h_1^{-1}(h_1, h_2, \cdots, h_s) \equiv (1, \tilde{h}_2, \cdots, \tilde{h}_s)(\widetilde{\text{mod}}\ n)$. 由于行置换不改变偏差值, 则两个好格子点相互等价. 因此, 可让 $h_1 = 1$, 即在生成向量中只需确定 $\{h_2, \cdots, h_s\}$. 因此, 搜索最佳好格子点时, 至多只需比较 $\begin{pmatrix} \phi(n)-1 \\ s-1 \end{pmatrix}$ 个好格子点.

(c) **Fibonacci 序列.** 对于二因子试验, 存在一种构造低偏差的格子点设计方法. 考虑 Fibonacci 序列 $1, 1, 2, 3, 5, 8, \cdots$, 定义 $F_0 = F_1 = 1$ 和 $F_m = F_{m-1} + F_{m-2}$, $m = 2, 3, \cdots$, 则当 $n = F_m$ 且 $\boldsymbol{h} = (1, F_{m-1})^{\mathrm{T}}$ 时, 其好格子点的偏差较小, 尤其当 $m \to \infty$ 时 (Hua and Wang, 1981, Section 4.8).

表 3.3 给出 $4 \leqslant n \leqslant 31$ 和 $s \leqslant 5$ 的最优生成向量, 其由 Fang 和 Ma (2001a) 第 98 页修改而得. 这里考虑由 (3.11) 式对因子个数的限制. 读者根据该表易得相应的好格子点.

表 3.3　在中心化偏差意义下的生成向量

n	$s=2$	$s=3$	$s=4$	$s=5$
4	(1 3)			
5	(1 2)	(1 2 3)		
6	(1 5)			
7	(1 3)	(1 2 3)	(1 2 3 5)	
8	(1 5)	(1 3 5)		
9	(1 4)	(1 4 7)	(1 2 4 7)	
10	(1 3)	(1 3 7)		
11	(1 7)	(1 5 7)	(1 2 5 7)	(1 2 3 5 7)
12	(1 5)	(1 5 7)		
13	(1 5)	(1 4 6)	(1 4 5 11)	(1 3 4 5 11)
14	(1 9)	(1 9 11)	(1 3 5 13)	
15	(1 11)	(1 4 7)	(1 4 7 13)	(1 2 4 7 13)
16	(1 7)	(1 5 9)	(1 5 9 13)	(1 3 5 9 13)
17	(1 10)	(1 4 10)	(1 4 5 14)	(1 4 10 14 15)
18	(1 7)	(1 7 13)	(1 5 7 13)	
19	(1 8)	(1 6 8)	(1 6 8 14)	(1 6 8 14 15)
20	(1 9)	(1 9 13)	(1 9 13 17)	(1 3 7 11 19)
21	(1 13)	(1 4 5)	(1 5 8 19)	(1 4 10 13 16)
22	(1 13)	(1 5 13)	(1 5 7 13)	(1 3 5 7 13)
23	(1 9)	(1 7 18)	(1 7 18 20)	(1 4 7 17 18)
24	(1 17)	(1 11 17)	(1 11 17 19)	(1 5 7 13 23)
25	(1 11)	(1 6 16)	(1 6 11 16)	(1 6 11 16 21)
26	(1 11)	(1 11 17)	(1 5 11 17)	(1 3 5 11 17)
27	(1 16)	(1 8 10)	(1 8 20 22)	(1 8 20 22 23)
28	(1 11)	(1 9 11)	(1 9 11 15)	(1 9 11 15 23)
29	(1 18)	(1 9 17)	(1 8 17 18)	(1 7 16 20 24)
30	(1 19)	(1 17 19)	(1 17 19 23)	(1 7 11 13 29)
31	(1 22)	(1 18 24)	(1 6 14 22)	(1 6 13 20 27)

3.3.2　删行好格子点法

在有些情形下, 例如 $s > k(n)$ 时, 候选集 $\mathcal{H}_n$ 中的元素已不足以产生列满秩的

好格子点. 然而, 若 $k(n+1)$ 大于 s 时, 可以考虑删行好格子点法来构造低偏差设计. 考虑下面的一个例子.

例 3.4 当 $n=12$ 时, 候选集 $\mathcal{H}_{12}=\{1,5,7,11\}$ 的元素个数为 4, 且由 (3.11) 可知 $k(12)=3$. 此时不能由好格子点法的算法3.1构造近似均匀设计 $U_{12}(12^s), s>3$. 然而, 由于

$$\mathcal{H}_{13}=\{1,2,3,4,5,6,7,8,9,10,11,12\}.$$

选择 $\mathcal{H}_{13}$ 中的 s 个元素可通过好格子点法获得 U-型设计 $U(13;13^s)$, 其设计的最后一行为 $(13,\cdots,13)$. 删除该行并保留其他 $12\times s$ 矩阵, 可得到一个 U-型设计 $U(12;12^s)$. 比较所有由这种方法得到的 12 行的设计, 可以得到近似均匀设计 $U_{12}(12^s)$. 这种推广的好格子点法被称为删行好格子点法 (leave one out good lattice point method).

删行好格子点法的具体过程如下所示:

算法 3.2 (删行好格子点法)

步骤 1 若 $s\leqslant\phi(n+1)/2$ 转步骤 2, 否则不能由该方法构造近似均匀设计.

步骤 2 给定 n 得到 $\mathcal{H}_{n+1}$, 记其元素为 $\boldsymbol{h}^*=(h_1^*,\cdots,h_m^*)$. 应用 (3.9) 的方法, 得到 $(n+1)\times m$ 设计矩阵 $\boldsymbol{U}(n+1,\boldsymbol{h}^*)$.

步骤 3 删除 $\boldsymbol{U}(n+1,\boldsymbol{h}^*)$ 的最后一行得到 $n\times m$ 矩阵 $\boldsymbol{H}$. 从 $\boldsymbol{H}$ 中选择可使得 U-型矩阵 $U(n;n^s)$ 为列满秩的 s 列. 记 $\mathcal{L}_{n,s}$ 为全体这类 U-型设计.

步骤 4 在 $\mathcal{L}_{n,s}$ 中选择具有最小偏差的设计, 得到近似均匀设计 $U_n(n^s)$.

易知, $\boldsymbol{U}(n+1,\boldsymbol{h}^*)$ 的第 $n+1$ 行为 $(n+1,\cdots,n+1)$. 删除这一行得到 $\boldsymbol{H}$, 其每一列都是 $(1,\cdots,n)^{\mathrm{T}}$ 的置换. 若 $\boldsymbol{H}$ 的第一行的两个元素 h_{1i} 和 h_{1j} 满足 $h_{1i}+h_{1j}=n+1$. 则 $\boldsymbol{H}$ 中的第 i 列 $\boldsymbol{h}_i$ 和第 j 列 $\boldsymbol{h}_j$ 之间为完全互相关, 因为 $\boldsymbol{h}_i+\boldsymbol{h}_j=(n+1)\mathbf{1}_n$. 因此, 若 $s>\phi(n+1)/2$, 则 $\boldsymbol{H}$ 中至少存在两列是完全互相关的. 具体细节可参见 (Yuan et al., 2017). 这也是算法 3.2 的步骤 1 中需要判断条件 $s\leqslant\phi(n+1)/2$ 是否成立的原因.

当 $\phi(n+1)>\phi(n)$ 时, 例如, n 为偶数时, 删行好格子点法的候选设计集 $\mathcal{L}_{n,s}$ 中的设计个数可能比好格子点法的候选设计集的设计个数多. 例如, 设 $n=36, s=4$. 好格子点法可能的候选设计集个数至多有 $\begin{pmatrix}\phi(36)-1\\4-1\end{pmatrix}=165$. 在中心化偏差意义下, 最好的好格子点的生成向量为 $(1,7,11,17)$, 其中心化偏差为 0.0029. 然而, 在删行好格子点法中, 可能的候选设计个数为 $\begin{pmatrix}\phi(36+1)-1\\4-1\end{pmatrix}=6545$, 远大于前者. 此时, 删行好格子点法得到的设计可能更均匀. 在本例中, 删行好格子点法的生成向量为 $(1,6,27,29)$, 相应的中心化偏差为 0.0023.

例 3.5 (例 3.4 续) 考虑用删行好格子点法构造均匀设计 $U_{12}(12^5)$. 从 3.3.1 节对于好格子点法的讨论可知, 生成向量的第一个元素可取为 1, 即 $h_1 = 1$. 可从剩余的 $\{2, 3, \cdots, 12\}$ 中选取 4 个元素, 则有 $\begin{pmatrix} 11 \\ 4 \end{pmatrix} = 330$ 个不同的生成向量. 其中生成向量 $(1, 2, 3, 4, 5)$ 在中心化偏差意义下是最优的. 矩阵 $\boldsymbol{H}$ 和相应的近似均匀设计 $U_{12}(12^5)$ 如下所示.

$\boldsymbol{H}$												$U_{12}(12^5)$				
1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5
2	4	6	8	10	12	1	3	5	7	9	11	2	4	6	8	10
3	6	9	12	2	5	8	11	1	4	7	10	3	6	9	12	2
4	8	12	3	7	11	2	6	10	1	5	9	4	8	12	3	7
5	10	2	7	12	4	9	1	6	11	3	8	5	10	2	7	12
6	12	5	11	4	10	3	9	2	8	1	7	6	12	5	11	4
7	1	8	2	9	3	10	4	11	5	12	6	7	1	8	2	9
8	3	11	6	1	9	4	12	7	2	10	5	8	3	11	6	1
9	5	1	10	6	2	11	7	3	12	8	4	9	5	1	10	6
10	7	4	1	11	8	5	2	12	9	6	3	10	7	4	1	11
11	9	7	5	3	1	12	10	8	6	4	2	11	9	7	5	3
12	11	10	9	8	7	6	5	4	3	2	1	12	11	10	9	8

表 3.4 给出 $n = 12$ 时的好格子点法和删行好格子点法的中心化偏差与相应的生成向量. 易知, 删行好格子点法往往比好格子点法更优, 即由删行好格子点法得到设计有更少的偏差值. 表 3.5 给出 $4 \leqslant n \leqslant 31$ 且 $s \leqslant 5$ 时的删行好格子点法的生成向量, 其由 Fang 和 Ma (2001a) 根据因子个数的限制而修改得到. Fang 和 Wang (1981)、Fang 和 Li (1995) 与 Fang 和 Ma (2001a) 证明对于许多 n 和 s, 由该方法得到的近似均匀设计比算法 3.1 的好格子点法得到的设计具有更低的偏差值.

表 3.4　$\boldsymbol{n = 12}$ 时的两种方法的中心化偏差和生成向量

构造方法	$s = 2$	$s = 3$	$s = 4$	$s = 5$
好格子点法	0.0506	0.1112		
	(1,5)	(1,5,7)		
删行好格子点法	0.0458	0.0782	0.1211	0.1656
	(1,5)	(1,3,4)	(1,2,3,5)	(1,2,3,4,5)

3.3.3 方幂好格子点法

即使对于一般大小的 n 和 s, 确定好格子点法的最佳生成向量也需要很多时间. Korobov (1959) 给出一种特殊形式的生成向量 $\boldsymbol{h} = (1, h, h^2, \cdots, h^{s-1})$ (mod

n), 其中 $\gcd(n,h^j)$=1 ($j=1,\cdots,s-1$) 且 $h<n$. 他证明对于合适的 h 和素数 n, 这种形式生成的设计的偏差会渐近小. 这种形式也被 Fang (1980)、Wang 和 Fang (1981) 与 Fang 和 Hickernell (1995) 用于搜索低偏差设计. 称这种方法为*方幂好格子点法*, 该方法的计算复杂度很低. 由这种方法得到的设计称为*方幂好格子点*. 给定正整数 (n,s), 该方法生成 (近似) 均匀设计的过程如下.

表 3.5 当 $4\leqslant n\leqslant 31$ 且 $s\leqslant 5$ 时, 删行好格子点法的生成向量

n	$s=2$	$s=3$	$s=4$	$s=5$
4	(1 2)			
6	(1 2)	(1 2 3)		
7	(1 3)			
8	(1 2)	(1 2 4)		
9	(1 3)	(1 3 7)	(1 3 7 9)	
10	(1 3)	(1 2 3)	(1 2 3 4)	(1 2 3 4 5)
11	(1 5)			
12	(1 5)	(1 3 4)	(1 2 3 5)	(1 2 3 4 5)
13	(1 3)	(1 3 5)		
14	(1 4)	(1 2 4)	(1 2 4 7)	
15	(1 7)	(1 3 5)	(1 3 5 7)	
16	(1 5)	(1 3 5)	(1 3 4 5)	(1 2 3 5 8)
17	(1 5)	(1 5 7)		
18	(1 7)	(1 7 8)	(1 3 4 5)	(1 2 5 6 8)
19	(1 9)	(1 3 7)	(1 3 7 9)	
20	(1 8)	(1 4 5)	(1 2 5 8)	(1 2 4 5 8)
21	(1 5)	(1 3 5)	(1 3 5 7)	(1 3 5 7 9)
22	(1 7)	(1 4 10)	(1 4 5 7)	(1 3 4 5 7)
23	(1 7)	(1 5 7)	(1 5 7 11)	
24	(1 7)	(1 4 11)	(1 4 6 9)	(1 4 6 9 11)
25	(1 7)	(1 3 7)	(1 3 5 7)	(1 3 5 7 9)
26	(1 8)	(1 8 10)	(1 4 5 7)	(1 2 5 7 8)
27	(1 5)	(1 3 5)	(1 3 5 11)	(1 3 5 9 11)
28	(1 12)	(1 8 12)	(1 8 9 12)	(1 4 5 7 13)
29	(1 11)	(1 7 11)	(1 7 11 13)	
30	(1 12)	(1 7 9)	(1 4 13 14)	(1 4 5 6 14)
31	(1 7)	(1 7 9)	(1 7 9 15)	(1 3 5 11 13)

算法 3.3 (方幂好格子点法)

步骤 1 确定正整数候选集

$$\mathcal{A}_{n,s}=\{a: a<n,\ \gcd(a,n)=1, \text{且}\ a,a^2,\cdots,a^s(\widetilde{\bmod}\ n)\ \text{互不相同}\}.$$

若集合 $\mathcal{A}_{n,s}$ 非空, 转步骤 2, 否则不能由该方法生成均匀设计.

步骤 2 对每个 $a \in \mathcal{A}_{n,s}$, 构造 U-型设计 $\boldsymbol{U}^a = (u_{ij}^a)$ 如下:

$$u_{ij}^a = ia^{j-1}\ (\widetilde{\text{mod}}\ n), \quad i = 1, \cdots, n,\ j = 1, \cdots, s.$$

步骤 3 确定 $a_* \in \mathcal{A}_{n,s}$, 使得 $\boldsymbol{U}^{a_*}$ 在所有的 $\boldsymbol{U}^a$ 中有最小的偏差值. 则 $\boldsymbol{U}^{a_*}$ 为近似均匀设计 $U_n(n^s)$.

对每个正整数 $a < n$, 若 n 和 a 互质, 则对于任意非负整数 i, n 和 a^i 也互质且 $\boldsymbol{U}^a$ 的任意列都是 $\{1, \cdots, n\}$ 的一个置换. 根据 n 和 a 互质的要求, $\mathcal{A}_{n,s}$ 中元素个数 $|\mathcal{A}_{n,s}| \leqslant \phi(n)$. 对于素数 n, 可以证明当 $s = n-1$ 时, $|\mathcal{A}_{n,s}| = \phi(\phi(n)) = \phi(n-1)$; 当 $s < n-1$ 时, $|\mathcal{A}_{n,s}| \in [\phi(n-1), n-1]$. 例如, $\phi(31) = 30$ 且 $\phi(30) = 8$. 则 $\mathcal{A}_{31,5}$ 的元素个数在 $[\phi(\phi(31)), 31-1] = [8, 30]$ 之间. 实际上, $|\mathcal{A}_{31,5}|$ =26. 在步骤 3 中, 当 $n = 31$ 时, 只需比较这 26 个 U-型设计. 许多均匀设计是由方幂好格子点法产生的, 具体参见 (Fang, 1980) 与 (Fang and Ma, 2001a). 删行的方法也可以应用到方幂好格子点法中. Fang 和 Li(1995) 证明由删行的方法得到的许多近似均匀设计比直接应用算法 3.3 得到的设计有更低的偏差.

Hua 和 Wang (1981) 证明了由好格子点法生成的 U-型设计的星偏差的收敛阶数为 $O(n^{-1}(\log\ n)^{s-1})$, 而由方幂好格子点法得到的星偏差的收敛阶数为

$$O(n^{-1}(\log\ n)^{s-1}\log(\log\ n)),$$

其比 $O(n^{-1}(\log\ n)^{s-1})$ 稍微差一点. 这从某种程度上说明了方幂好格子点法的性能也是不错的.

3.4 切 割 法

当 n 不是素数时, 考虑构造相应的均匀设计 $U_n(n^s)$. 对于较小的 s, 也可以考虑用好格子点法构造设计, 不过其均匀性并不是太好; 对于较大的 s, 好格子点法可能会失效, 因为好格子点法至多能构造 $\phi(n)/2+1$ 列.

对于任意的 n, Ma 和 Fang(2004) 建议用**切割法**来构造 $U_n(n^s)$. 切割法的主要思想为: 设 $\boldsymbol{U}_p$ 为一个均匀设计 $U_p(p^s)$, 其中 $n < p$ 或 $n \ll p$ 且 p 或 $p+1$ 为素数. 记 $\mathcal{P}_p$ 为其导出设计. 设 $\mathcal{Q}$ 为 $[0,1]^u$ 的子区域使得 $\mathcal{P}_p$ 落入 $\mathcal{Q}$ 的个数恰好为 n 个, 并记 $\mathcal{P}_n$ 为这 n 个点. 由伪蒙特卡罗法的理论可知, $\mathcal{P}_n$ 也均匀地散布在区域 $\mathcal{Q}$ 中. 特别地, 可以选择 C^s 中合适的超矩形使得 $\mathcal{P}_p$ 中恰好有 n 个点落在该超矩形中. 通过线性变换, 可以把这 n 个点变成一个近似均匀设计. 图 3.3 给出一个直观的例子, 其中 $p = 47$, $n = 8$ 和 $s = 2$. 选择合适的超矩形使得落入其中的个数恰好为 $n = 8$; 然后重设这些试验点的水平数, 获得一个 n 水平的 U-型设计. 我们可以

移动超矩形并寻找一个均匀性最优的设计. 注意这个超矩形可以有可卷性质. 实际上, 从 $\boldsymbol{U}_p$ 中直接获得 n 个点近似均匀设计的切割法的具体过程如下所示.

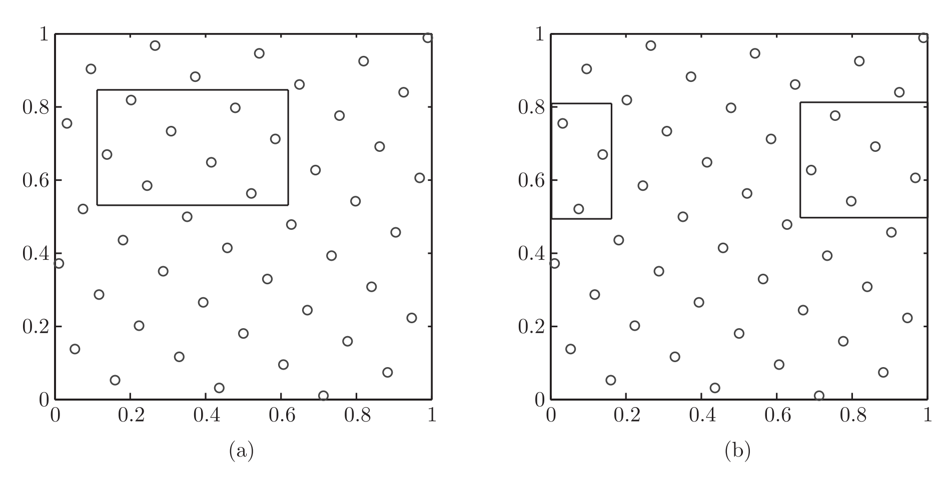


图 3.3　切割法示例

算法 3.4 (切割法)

步骤 1　**原始设计**　给定 (n,s), 寻找均匀设计 $U_p(p^s)$, 其中 $p \gg n$ 且 p 或 $p+1$ 为素数, 记之为 $\boldsymbol{U}_p=(u_{ij})$, 并称为初始设计.

步骤 2　**行排列**　对于 $l=1,\cdots,s$, 根据第 l 列, 置换 $\boldsymbol{U}_p$ 的行序, 使得该列的元素恰好从 1 到 p 排序. 记行置换后的矩阵为 $\boldsymbol{U}_p^{(l)}=(u_{kj}^{(l)})$.

步骤 3　**切割**　对于 $m=1,\cdots,p$, 令 $\boldsymbol{U}_p^{(l,m)}=(u_{kj}^{(l,m)})$, 其中

$$u_{kj}^{(l,m)}=\begin{cases} u_{k+m-n-1,j}^{(l)}, & m>n,\ k=1,\cdots,n, \\ u_{kj}^{(l)}, & m\leqslant n,\ k=1,\cdots,m-1, \quad j=1,\cdots,s. \\ u_{k+p-n,j}^{(l)}, & m\leqslant n,\ k=m,\cdots,n, \end{cases}$$

步骤 4　**新设计空间**　把 $\boldsymbol{U}_p^{(l,m)}$ 中每列的元素按照从小到大重新标记为 1, $2,\cdots,n$. 则得到的矩阵变成一个 U-型矩阵 $U(n;n^s)$, 并记为 $\boldsymbol{U}^{(l,m)}$. 我们一共有 ps 个这样的 U-型矩阵.

步骤 5　**输出设计**　给定均匀性测度 D, 比较步骤 4 得到的这 ps 个设计 $\boldsymbol{U}^{(l,m)}$, 并选择一个具有最小 D-值的设计. 这即近似均匀设计 $U_n(n^s)$.

在步骤 1 中, 若好格子点法的复杂度可承受, 则可用好格子点法获得一个好的初始设计；否则, 可考虑方幂好格子点法, 因为其是获得较好的初始设计的最快速方法. 下面给出一个例子来说明.

例 3.6 (构造 $U_8(8^3)$) 考虑由好格子点法构造的近似均匀设计 $U_{47}(47^3)$, 其生成向量为 $(1,18,26)$, 初始设计为

$$
\boldsymbol{U}_{47}=\left(\begin{matrix}
1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16&17&18&19&20&21&22&23&24\\
18&36&7&25&43&14&32&3&21&39&10&28&46&17&35&6&24&42&13&31&2&20&38&9\\
26&5&31&10&36&15&41&20&46&25&4&30&9&35&14&40&19&45&24&3&29&8&34&13
\end{matrix}\right.
$$
$$
\left.\begin{matrix}
25&26&27&28&29&30&31&32&33&34&35&36&37&38&39&40&41&42&43&44&45&46&47\\
27&45&16&34&5&23&41&12&30&1&19&37&8&26&44&15&33&4&22&40&11&29&47\\
39&18&44&23&2&28&7&33&12&38&17&43&22&1&27&6&32&11&37&16&42&21&47
\end{matrix}\right)^{\mathrm{T}}.
$$

$\boldsymbol{U}_{47}$ 的中心化偏差为 6.57×10^{-4}. 考虑从 $\boldsymbol{U}_{47}$ 中构造 $U_8(8^3)$. 可以根据行置换后的第 1, 第 2 或第 3 列的某一行开始, 依次选择 8 行. 对于每一列, 有 47 种选择, 则总共有 $141=47\times3$ 种选择, 即 $\boldsymbol{U}_p^{(l,m)}$, $l=1,2,3$, $m=1,\cdots,47$. 重标记水平后, 都可以得到一个 U-型设计, 其中具有最小中心化偏差的是 $\boldsymbol{U}^{(3,9)}$, 即 $l=3, m=9$. 首先根据第 3 列行置换初始设计, 有

$$
\boldsymbol{U}_{47}^{(3)}=\left[\begin{matrix}
38&29&20&11&2&40&31&22&13&4&42&33&24&15&6&44&35&26&17&8&46&37&28&19\\
26&5&31&10&36&15&41&20&46&25&4&30&9&35&14&40&19&45&24&3&29&8&34&13\\
1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16&17&18&19&20&21&22&23&24
\end{matrix}\right.
$$
$$
\left.\begin{matrix}
10&1&39&30&21&12&3&41&32&23&14&5&43&34&25&16&7&45&36&27&18&9&47\\
39&18&44&23&2&28&7&33&12&38&17&43&22&1&27&6&32&11&37&16&42&21&47\\
25&26&27&28&29&30&31&32&33&34&35&36&37&38&39&40&41&42&43&44&45&46&47
\end{matrix}\right]^{\mathrm{T}}.
$$

则

$$
\boldsymbol{U}_p^{(3,9)}=\begin{bmatrix}
38&29&20&11&2&40&31&22\\
26&5&31&10&36&15&41&20\\
1&2&3&4&5&6&7&8
\end{bmatrix}^{\mathrm{T}},
$$

并从小到大重标记水平可得

$$
\boldsymbol{U}^{(3,9)}=\begin{bmatrix}
7&5&3&2&1&8&6&4\\
5&1&6&2&7&3&8&4\\
1&2&3&4&5&6&7&8
\end{bmatrix}^{\mathrm{T}},
$$

其为近似均匀设计 $U_8(8^3)$, 其中心化偏差为 0.0129. 相比较而言, 当 $n=8, s=3$ 时, 好格子点法得到的设计的中心化偏差为 0.0132. 因此, 由切割法得到的 $U_8(8^3)$ 比好格子点法得到的更优.

切割法有几个优点: ①给定一个初始设计 $U_p(p^s)$, 如果 p 足够大, 可以得到许多近似均匀设计 $U_n(n^s)$, $n<p$; ② 由切割法得到的设计通常比好格子点法得到的设计有更好的均匀性; ③ 切割法的性能不依赖于特定的均匀性测度. 更多的细节和例子可参见 (Ma and Fang, 2004).

3.5 线性水平置换法

给定参数 (n,s), 由好格子点法生成的设计空间是 U-型设计空间 $\mathcal{U}(n;n^s)$ 的子集. 我们可以找到一些方法改进好格子点的均匀性. Tang 等 (2012) 证明水平置换方法可以改进正规设计的空间填充性. Zhou 和 Xu (2015) 证明好格子点经线性水平置换之后也可以提高其空间填充性. 作为一类特殊的均匀设计, 好格子点也可以应用到计算机试验中, 其逐渐变为研究复杂系统的常见方法 (Sacks et al., 1989a; Santner et al., 2003; Fang et al., 2006a). 计算机试验需要大设计矩阵. 例如, Morris(1991) 考虑一个有几百个因子的仿真实模型, 此时一个小的设计矩阵不足以刻画因子与响应之间的关系. 好格子点可用于构造大设计. 本节考虑线性水平置换技术用于提高好格子点在均匀性以及极大极小距离意义下的空间填充性.

Johnson 等 (1990) 提出极大极小距离和极小极大距离用于衡量设计的空间填充性. 对任意设计 $\boldsymbol{U}\in\mathcal{D}(n;q^s)$, 记其 q 个水平为 $0,1,\cdots,q-1$. 定义 $d_p(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^s|x_i-y_i|^p, p\geqslant 1$ 为 $\boldsymbol{U}$ 中任意两行 $\boldsymbol{x}=(x_1,\cdots,x_s)$ 和 $\boldsymbol{y}=(y_1,\cdots,y_s)$ 的L_p-距离. 当 $p=1$ 时, 该距离为*矩形距离*, 当 $p>1$ 时, L_p-距离即 L_p-范数的 p 次幂. 设计 $\boldsymbol{U}$ 的 L_p-距离定义为

$$d_p(\boldsymbol{U})=\min\{d_p(\boldsymbol{x},\boldsymbol{y}):\boldsymbol{x}\neq\boldsymbol{y},\boldsymbol{x},\boldsymbol{y}\ \text{为}\ \boldsymbol{U}\ \text{的行}\}. \tag{3.12}$$

(Johnson et al., 1990) 中的极大极小距离准则即为 $\max d_p(\boldsymbol{U})$.

定义 3.3 对于正整数 $q\geqslant 2$, 称一个 s 个因子、q 水平且有 q^{s-k} 个试验点的设计 $\boldsymbol{U}$ 为正规 q^{s-k} 设计, 若 $\boldsymbol{U}$ 由一个定义在 $Z_q=\{0,1,\cdots,q-1\}$ 上完全因子设计 $q^{(s-k)}$, M 和一个生成矩阵 $\boldsymbol{G}$ 的乘积而成, 即

$$\boldsymbol{U}=\boldsymbol{M}\boldsymbol{G}\pmod q. \tag{3.13}$$

这里用这种方式来定义正规 q^{s-k} 设计的目的是考虑正规设计不同行之间的距

离. 例如, 考虑 $q=3, s=4, k=2$ 的情形, $\boldsymbol{M}$, $\boldsymbol{G}$ 和 $\boldsymbol{U}$ 如下所示

$$\boldsymbol{M}=\begin{bmatrix}0&0\\0&1\\0&2\\1&0\\1&1\\1&2\\2&0\\2&1\\2&2\end{bmatrix},\quad \boldsymbol{G}=\begin{bmatrix}1&0&1&1\\0&1&1&2\end{bmatrix},\quad \boldsymbol{U}=\begin{bmatrix}0&0&0&0\\0&1&1&2\\0&2&2&1\\1&0&1&1\\1&1&2&0\\1&2&0&2\\2&0&2&2\\2&1&0&1\\2&2&1&0\end{bmatrix}.$$

正如 (Tang et al., 2012) 与 (Zhou and Xu, 2014) 所示, 置换一个或多个因子的水平并不改变设计的正交性, 但是可以改变其几何结构和空间填充性. 对于正规设计, 考虑其在 $Z_q=\{0,1,\cdots,q-1\}$ 上的线性水平置换, 其有一个简单的形式. 给定一个正规 q^{s-k} 设计 $\boldsymbol{U}$ 和一个 $s\times 1$ 的向量 $\boldsymbol{u}$, 记

$$\boldsymbol{U}_{\boldsymbol{u}}=\boldsymbol{U}+\boldsymbol{u}=\{\boldsymbol{x}+\boldsymbol{u}\ (\bmod\ q):\boldsymbol{x}\ \ 为\ \boldsymbol{U}的每一行\} \tag{3.14}$$

为 $\boldsymbol{U}$ 在 Z_q 上经线性水平置换后的设计, 其中

$$\boldsymbol{x}+\boldsymbol{u}\ (\bmod\ q)=\{x_1+u_1,\cdots,x_s+u_s\}\ (\bmod\ q).$$

当考虑所有可能的线性水平置换, 只需考虑置换 k 个相关列, 而保持 $(s-k)$ 列独立列不变. 这里独立列指的是没有一列可以由其他列在模 q 意义下的线性组合而得到的. 不失一般性, 可假设正规 q^{s-k} 设计的前面 $(s-k)$ 列为独立列, 即前面 $(s-k)$ 列恰好构成一个完全因子设计. 对于正规设计的 L_p-距离, 有下面的结论.

定理 3.2 对于定义在 Z_q 上的正规 q^{s-k} 设计 $\boldsymbol{U}$, 任意线性水平置换都不会减少其 L_p-距离, 即 $d_p(\boldsymbol{U}_{\boldsymbol{u}})\geqslant d_p(\boldsymbol{U})$, 其中 $\boldsymbol{U}_{\boldsymbol{u}}$ 如 (3.14) 式定义.

定理 3.2 的证明见 (Zhou and Xu, 2015). 定理 3.2 说明了在极大极小距离意义下线性水平置换有可能正规设计的提高空间填充性. 不过需要指出的是, 存在一些设计, 对其做线性水平置换既不增加又不减少 L_p-距离. 例如, 可以证明对于任意正规 3^{s-1} 设计, 任意线性水平置换不改变设计的 L_p-距离.

根据正规设计的定义 (3.13), 好格子点 $\boldsymbol{U}$ 可以由一个正规设计变化而来, 即把其中的水平 0 变为 n 即可. 特别地, 设 $\boldsymbol{U}_0$ 为一个生成矩阵为 $\boldsymbol{h}$ 的正规设计, 其独立列为 $M=(0,1,\cdots,n-1)^{\mathrm{T}}$, 即

$$\boldsymbol{U}_0=M\boldsymbol{h}\quad (\bmod\ n). \tag{3.15}$$

通过把 $\boldsymbol{U}_0$ 中第一行的 $(0,\cdots,0)$ 变换为 $(n,\cdots,n)$, 可获得一个好格子点 $\boldsymbol{U}$. 因此, 好格子点可被认为是一种特殊的正规设计. 好格子点的元素都来自于集合 $\{1,\cdots,n\}$, 而定义 3.3 中正规设计的元素都来自于集合 $\{0,1,\cdots,n-1\}$. 在 1.2 节中给出更多说明.

把 0 置换为 n, 设计的几何结构和统计性质会有一些改变, 但下面的结果说明对于特殊的正规设计而言这种变化不改变其 L_p-距离.

推论 3.3　由 (3.15) 定义的正规设计 $\boldsymbol{U}_0$ 与相应的好格子点 $\boldsymbol{U}$ 有同样的 L_p-距离.

证明　对于任意非零行 $\boldsymbol{x}=(x_1,\cdots,x_n)\in\boldsymbol{U}_0$, 存在一个整数 $c\in\{1,\cdots,n-1\}$ 使得 $\boldsymbol{x}=c\boldsymbol{h} \pmod n$. 则 $\boldsymbol{z}=(n-c)\boldsymbol{h}=(n-x_1,\cdots,n-x_n) \pmod n$ 为 $\boldsymbol{U}$ 的某一行. 则 $\boldsymbol{x}$ 和行向量 $(0,\cdots,0)$ 之间的 L_p-距离恰好等于 $\boldsymbol{z}$ 和行向量 $(n,\cdots,n)$ 之间的 L_p-距离. 由于 $\boldsymbol{U}$ 和 $\boldsymbol{U}_0$ 中 $n-1$ 个设计点都是一样的, 且这 $n-1$ 个点和 $(0,\cdots,0)$ 之间的 $n-1$ 个 L_p-距离的分布与其和 $(n,\cdots,n)$ 之间的 $n-1$ 个 L_p-距离的分布是一致的, 因此把 $\boldsymbol{U}_0$ 中的 0 置换为 n 并不影响整个设计的 L_p- 距离. 定理得证.

考虑对好格子点进行线性水平置换 (3.14), 置换后用 n 替换 0. 由定理 3.2 可知, 任意线性水平置换都不减少其 L_p-距离. 类似于定理 3.2, 对于好格子点, 有下面的结果.

命题 3.4　设 $\boldsymbol{U}$ 为 $n\times s$ 的好格子点, 其任意线性水平置换都不减少其 L_p-距离, 即 $d_p(\boldsymbol{U}_{\boldsymbol{u}})\geqslant d_p(\boldsymbol{U})$, 其中 $\boldsymbol{U}_{\boldsymbol{u}}$ 如 (3.14) 定义, 并用 n 替换 0.

由命题 3.4 可知, 对好格子点做线性水平置换可以提高其空间填充性. 为此, 把线性水平置换后的设计称为*广义好格子点*. 由前面的讨论可知, 构造广义好格子点的算法如下所示.

算法 3.5 (构造广义好格子点)

步骤 1　给定设计准则和参数 n 和 s, 由好格子点法获得初始设计 D_0.

步骤 2　选择一个置换向量 $\boldsymbol{u}$, 获得设计 $D_{\boldsymbol{u}}$ 及相应的准则值 $f(\boldsymbol{u})$.

步骤 3　若 $f(\boldsymbol{u})$ 已达到其上 (下) 界, 则 $D_{\boldsymbol{u}}$ 即为最终的设计并终止算法; 否则转步骤 2 直到重复次数达到 K.

步骤 4　在这 K 个设计中, 获得在给定准则下的最佳设计 D^*, 其即为广义好格子点.

算法 3.5 的步骤 1 中预先给定的准则可以是某个均匀性度量, 如中心化偏差、可卷偏差或混合偏差等, 也可以是极大极小距离准则. 对于均匀性度量, 步骤 3 中选取为下界, 其可参考 2.6 节; 当考虑极大极小距离准则时, 可以证明对于任意设计 $D\in\mathcal{U}(n;n^s)$, 其 L_1-距离的上界为 $\lfloor(n+1)s/3\rfloor$, 而 L_2-距离的上界 $\lfloor n(n+1)s/6\rfloor$,

其中 $\lfloor\cdot\rfloor$ 表示向下取整. 当 n 和 s 比较小时, K 可以取为线性水平置换的所有可能数, 即 $K=n^{s-1}$, 其只需考虑对好格子点的后 $s-1$ 个相关列做线性水平置换. 否则 K 可取为计算机能承受的值, 例如, $K=10^4\sim10^6$.

例 3.7 设 $n=7, s=6$. 考虑生成向量 $\boldsymbol{h}=(1,\cdots,6)$, 得到好格子点

$$\boldsymbol{U}=\begin{bmatrix}1&2&3&4&5&6\\2&4&6&1&3&5\\3&6&2&5&1&4\\4&1&5&2&6&3\\5&3&1&6&4&2\\6&5&4&3&2&1\\7&7&7&7&7&7\end{bmatrix},$$

其中 $d_1(\boldsymbol{U})=12$. 不失一般性, 考虑对 U 的最后 5 个相关列做的所有线性水平置换. 在所有的 16807 个置换设计, 有 16167 个设计的 L_1-距离与 $\boldsymbol{U}$ 相同, 而有 640 个设计的 L_1-距离为 13. 例如, 考虑置换向量 $\boldsymbol{u}=(0,4,1,5,2,6)$, 可得设计

$$\boldsymbol{U_u}=\begin{bmatrix}1&6&4&2&7&5\\2&1&7&6&5&4\\3&3&3&3&3&3\\4&5&6&7&1&2\\5&7&2&4&6&1\\6&2&5&1&4&7\\7&4&1&5&2&6\end{bmatrix},$$

其中 $d_1(\boldsymbol{U_u})=13$. 注意 $U_{\boldsymbol{u}}$ 等价于 $\boldsymbol{U}+3\ (\widetilde{\mathrm{mod}}\ 7)$, 其有相同的设计点. 实际上 $\boldsymbol{U}+5\ (\widetilde{\mathrm{mod}}\ 7)$ 也有同样的 L_1-距离 13.

命题 3.4 说明线性水平置换可以增加其 L_1-距离. 一般地, 增加极大极小距离的同时也会改善均匀性, 即偏差也会减少. 然而寻找一个最优的线性水平置换并不容易, 尤其是 n 和 s 都比较大的情形, 例 3.7 说明也可以关注一些特殊的线性水平置换. 例如, 考虑简单线性水平置换 $\boldsymbol{U}+i\ (\mathrm{mod}\ n), i=1,\cdots,n-1$, 其往往也能找到更好的空间填充设计.

例 3.8 设 $n=37$, $s=36$. 则相应唯一的好格子点的 L_1-距离和中心化偏差分别为 342 和 1200.89. 考虑简单线性水平置换 $\boldsymbol{U}+i\ (\widetilde{\mathrm{mod}}\ n)$, $i=1,\cdots,n-1$. 在这 36 个置换设计中, 记 $\boldsymbol{U}_1^*$ 和 $\boldsymbol{U}_2^*$ 分别为具有最大 L_1-距离的设计和最小偏差的设计. 易知, U_1^* 的 L_1-距离和中心化偏差分别为 408 和 203.9, U_2^* 的 L_1-距离和中心化偏差分别为 342 和 61.48. 这说明中心化偏差可以被显著减少, 且简单线性水

平置换可以提高其空间填充性. 另一方面, 可以随机产生 $K=10^4$ 个置换向量, 并得到 K 个置换设计. 在这 K 个设计中, U_1^* 的 L_1-距离和中心化偏差分别为 376 和 79.15, U_2^* 的 L_1-距离和中心化偏差分别为 358 和 63.34. 综上所述,

$$
\begin{aligned}
&D_p(\boldsymbol{U})=342, \quad \mathrm{CD}(\boldsymbol{U})=1200.89,\\
&D_p(\boldsymbol{U}_1^*)=408, \quad \mathrm{CD}(\boldsymbol{U}_1^*)=203.9,\\
&D_p(\boldsymbol{U}_2^*)=342, \quad \mathrm{CD}(\boldsymbol{U}_2^*)=61.48,\\
&D_p(\boldsymbol{U}_1^*)=376, \quad \mathrm{CD}(\boldsymbol{U}_1^*)=79.15, \text{ 在 } K \text{ 个设计中},\\
&D_p(\boldsymbol{U}_2^*)=358, \quad \mathrm{CD}(\boldsymbol{U}_2^*)=63.34, \text{ 在 } K \text{ 个设计中}.
\end{aligned}
$$

显然, 在均匀性和极大极小距离意义下, 空间填充性都有所提高.

例 3.9　设 $n=89$, $s=24$. 初始设计 $\boldsymbol{U}$ 由方幂好格子点法产生, 其生成元为 $a=33$, 相应设计的中心化偏差为 3.4302. 考虑置换向量

$$\boldsymbol{u}=(0\ 20\ 25\ 64\ 45\ 7\ 82\ 8\ 9\ 31\ 87\ 1\ 12\ 62\ 13\ 49\ 37\ 10\ 55\ 33\ 32\ 52\ 41\ 81),$$

可以获得设计 $\boldsymbol{U}_{\boldsymbol{u}}$, 其中心化偏差为 1.7146, 这比 $\boldsymbol{U}$ 的中心化偏差小. 因此线性水平置换可以提高均匀性.

例 3.7— 例 3.9 说明简单的线性水平置换 $\boldsymbol{U}+i \pmod n$ 是有用的. 进一步地, 随着置换向量数 K 的增大, 我们往往可以得到空间填充性更好的设计.

好格子点容易构造, 且我们证明线性水平置换可以增加其极大极小距离及减少其偏差. 一个问题是如何在极大极小距离意义下, 找到一个好的或最佳的线性水平置换. 对于一个小设计, 可以比较所有的线性水平置换. 然而, 当行数和因子数比较大时, 不能完全搜索. 一种简单的方法是随机产生一些置换向量, 并比较其置换设计, 选取其中最佳的设计. 另一种方法是考虑如模拟退火算法、门限接受法等随机搜索算法, 确定一个最佳的线性水平置换. 具体方法见第 4 章.

3.6　组合构造方法

本节主要采用 2.5.1 小节介绍的离散偏差来度量设计的均匀性. 如果一个设计达到离散偏差的下界, 那么该设计就是在离散偏差准则下的均匀设计. 关于离散偏差的一些紧下界已经在 2.6.4 小节给出, 因此, 我们需要做的就是发展一些理论方法来构造达到这些偏差下界的设计.

3.6.1 小节和 3.6.2 小节关注如何通过组合设计来构造均匀设计, 将给出均匀设计与均匀可分解设计的联系以及相应的几种构造均匀设计的方法. 3.6.3 小节讨论基于饱和正交表的构造方法. 3.6.4 小节介绍某些进一步的结果, 包含投影均匀性和近似 U-型设计的结果.

3.6.1 均匀设计与均匀可分解设计的联系

组合设计在统计试验设计和纠错码理论中有广泛应用. 而且, 近年来, 其应用更是遍及实验和理论计算科学、通信、密码学和网络等各个领域. 关于这方面的综述, 读者可以参考 (Colbourn et al., 1999). 在本小节及 3.6.2 小节, 我们关注组合设计理论在试验设计理论中的一个新应用, 即由均匀可分解不完全区组设计构造均匀设计. 首先, 介绍均匀可分解设计的一些相关知识, 并说明它们与均匀设计的关系.

区组 (block) 是具有十分相近的试验环境的集合. 比如, 某化学过程中的一个试验可能需要两台仪器. 仪器之间可能存在系统性差异. 如果我们对两台仪器的效应不感兴趣, 那么仪器可被视为讨厌变量. 在这个试验中, 由于同一台仪器下试验环境的变异性较之不同仪器下的变异性更小, 故每台仪器构成一个区组.

例 3.10 在某医学试验中, 试验者想要比较五种药物. 每种药物需要 10 只小鼠, 一共需要 50 只小鼠. 在该试验中, 应考虑小鼠的变异性. 由于来自同一家族的小鼠间变异性较小, 故最为理想的情况是 50 只小鼠都来自同一家族. 但实际很难做到. 取而代之的是, 可以考虑选取 10 个家族, 并从每个家族选取 5 只小鼠. 在这里, 来自同一家族的 5 只小鼠构成一个区组. 整个试验中共有 10 个区组.

定义 3.4 (区组设计) 假设将 n 个处理安排到 b 个区组, 记第 j 个区组有 k_j 个试验单元, 第 i 个处理在整个设计中出现 s_i 次, 这里 $i=1,\cdots,n, j=1,\cdots,b$.

1) 如果一个区组设计不是所有处理都出现在每个区组中, 则称该区组设计为不完全的 (incomplete).

2) 如果一个区组设计对所有的 i 有 $s_i=s$, 则称之为等重复的 (equireplicate).

3) 如果一个区组设计对所有的 j 有 $k_j=k$, 则称之为相宜的 (proper).

4) 如果一个区组设计的每个处理在每个区组中不出现或者仅出现一次, 则称之为二元的 (binary).

注意到, 在之前的章节中 n 和 s 用来表示的是试验次数和因子数, 而此处它们分别表示处理和区组的个数, 从接下来的讨论中可以看到两者是一致的. 在区组设计中, 等重复的、相宜的二元不完全区组设计受到广泛关注, 其中, 应用最广泛的是平衡不完全区组设计 (balanced incomplete block design, BIBD).

定义 3.5 如果一个等重复的、相宜的二元不完全区组设计满足任一处理对都同时出现在 λ 个区组中, 则称之为平衡不完全区组设计, 记作 BIBD(n,b,s,k,λ).

显然这五个参数满足下面两个关系式:

$$ns=bk \quad \text{且} \quad \lambda(n-1)=s(k-1). \tag{3.16}$$

因此, 可以只用三个参数 n,k,λ 来标记 BIBD, 即 BIBD(n,k,λ).

定义 3.6 对于一个区组设计, 如果它的区组可以分为若干个平行类, 使得每

个处理在每个平行类中都出现且只出现一次, 则称之为可分解的 (resolvable). 一个平行类, 如果其中的每个区组大小相同, 则称之为均匀的 .

如果一个区组设计是可分解的, 显然它也是等重复的和二元的. 一个可分解的 BIBD(n,k,λ) 记作 RBIBD(n,k,λ).

记 $\mathcal{A}$ 为区组大小的不同值构成的集合, $\mathcal{R}$ 为多重集合, 满足 $|\mathcal{R}|=|\mathcal{A}|$, 其中 $|\mathcal{A}|$ 为集合 $\mathcal{A}$ 的基数.

定义 3.7 假设对于每个 $k\in\mathcal{A}$, 对应有一个正的 $r_k\in\mathcal{R}$, 相应地有 r_k 个区组大小都为 k 的平行类. 用 URBD$(n,\mathcal{A},\mathcal{R})$ 去表示一个由均匀平行类构成的可分解不完全区组设计. 一个 URBD$(n,\mathcal{A},\mathcal{R})$, 如果任一处理对都恰出现在 λ 个区组中, 则称之为均匀可分解设计 (uniformly resolvable design, URD), 记作 URD$(n,\mathcal{A},\lambda,\mathcal{R})$.

对一个 URD, 记 $\mathcal{A}=\{k_1,\cdots,k_l\}$, $\mathcal{R}=\{r_1,\cdots,r_l\}$, 显然有

$$\lambda(n-1)=\sum_{i=1}^{l} r_i(k_i-1). \tag{3.17}$$

关于区组设计及设计理论的一般背景的详细讨论, 可以参见(Caliński and Kageyama, 2000) 及 (Beth et al., 1999). 为了内容的完整性并方便阅读, 下面给出与 (Fang et al., 2006a) 的 3.5 节一些类似的例子及其讨论. 先来看一个 URD 的例子.

例 3.11 (一个 URD$(6,\{3,2\},1,\{1,3\})$) 假设将 $n=6$ 个处理安排到 $b=11$ 个区组中, 区组大小 k_j 来自集合 $\mathcal{A}=\{3,2\}$(即 $k_1=3,k_2=2$), 使得每个处理恰好在四个平行类中出现, 记为 P_1,P_2,P_3 和 P_4, 如下所示 (在每个平行类中, 一个 $\{\cdots\}$ 代表一个区组).

$$P_1=\{\{1,2,3\},\{4,5,6\}\};$$
$$P_2=\{\{1,4\},\{2,5\},\{3,6\}\};$$
$$P_3=\{\{3,5\},\{1,6\},\{2,4\}\};$$
$$P_4=\{\{2,6\},\{3,4\},\{1,5\}\}.$$

注意任一处理对都出现在 $\lambda=1$ 个区组中, 因此它是 URD$(6,\{3,2\},1,\{1,3\})$.

下面讨论 U-型设计与 URD 设计间的关系. 假设有一个 URD$(n,\mathcal{A},\lambda,\mathcal{R})$. 为方便起见, 记 n 个处理为 $1,\cdots,n$, 令 $\mathcal{A}=\{n/q_1,\cdots,n/q_s\}$, $\mathcal{R}=\{1,\cdots,1\}$, 并记区组大小为 n/q_j 的平行类为 P_j, $j=1,\cdots,s$. 那么可以按照下面的算法 3.6 构造一个 U-型设计 $\boldsymbol{U}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$.

算法 3.6 (URD–UD)

步骤 1 给每一个平行类 P_j 中的 q_j 个区组一个自然序 $1,\cdots,q_j$, 其中 $j=1,\cdots,s$.

步骤 2 对每一个 P_j, 用如下方法构造一个 q_j-水平的列 $\boldsymbol{u}^j=(u_{ij})$: 如果处理 i 包含在 P_j 的第 u 个区组中, 则取 $u_{ij}=u$, $u=1,\cdots,q_j$.

步骤 3 由 $P_j(j=1,\cdots,s)$ 构造的 s 个列构成一个 U-型设计 $\boldsymbol{U}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$.

从上述方法可以看出, 对 $\boldsymbol{U}$ 的任意两不同行 $\boldsymbol{u}_i$ 和 $\boldsymbol{u}_j$ 而言, 它们的相遇数恰好是同时包含处理 i 和 j 的区组的个数, 为常数 λ. 由 (3.17) 可知 $\lambda=\left(\sum_{j=1}^{s}\dfrac{n}{q_j}-s\right)\Big/(n-1)$. 根据定理 2.22, 所构造的设计 $\boldsymbol{U}$ 是在离散偏差准则下的一个均匀设计 $U_n(q_1\times\cdots\times q_s)$. 定理 3.5 对上述阐述进行了总结.

下面取例 3.11 中给出的 $\mathrm{URD}(6,\{3,2\},1,\{1,3\})$ 为例做一下说明.

例 3.12 (一个 $U_6(2^13^3)$) *由例 3.11 给出的 $\mathrm{URD}(6,\{3,2\},1,\{1,3\})$ 的四个平行类, 可通过下面的方法构造出 $\boldsymbol{U}$ 的四列. 在 P_1 中, 有两个区组 $\{1,2,3\}$ 和 $\{4,5,6\}$. 把 “1” 放到 $\boldsymbol{U}$ 的第一列的第 1, 2, 3 行, 并且将 “2” 放到这一列的第 4, 5, 6 行, 这样得到 $\boldsymbol{U}$ 的一个二水平列. 类似地, 对 P_2, 把 “1” 放到 $\boldsymbol{U}$ 的第二列的第 1, 4 行, 并且将 “2” 和 “3” 分别放到该列的第 2, 5 行及第 3, 6 行. 这样得到 $\boldsymbol{U}$ 的一个三水平列. 以这种方式, 由 4 个平行类可以构造出 4 列, 构成如下 U-型设计 $\boldsymbol{U}$:*

$$\boldsymbol{U}=\begin{bmatrix}1&1&2&3\\1&2&3&1\\1&3&1&2\\2&1&3&2\\2&2&1&3\\2&3&2&1\end{bmatrix},$$

显然 $\boldsymbol{U}$ 是一个均匀设计 $U_6(2^13^3)$.

反过来, 给定一个均匀设计, 若其任意两行的相遇数为常数, 则可以由该均匀设计构造一个 URD. 设 $\boldsymbol{U}$ 是这样一个均匀设计 $U_n(q_1\times\cdots\times q_s)$, 构造方法如下:

算法 3.7 步骤 1 对于 $\boldsymbol{U}$ 的每一列 $\boldsymbol{u}^j$, 用如下方法构造一个由 q_j 个大小为 n/q_j 的不相交区组构成的平行类: 如果 $\boldsymbol{u}^j$ 的第 i 个元素取 u 水平, 则让第 u 个区组包含处理 i, $i=1,\cdots,n$; $u=1,\cdots,q_j$.

步骤 2 对 $j=1,\cdots,s$, 将由步骤 1 构造的 s 个平行类合到一起得到一个 $\mathrm{URBD}(n,\mathcal{A},\mathcal{R})$, 其中 $\mathcal{A}=\{n/q_1,\cdots,n/q_s\}$ 且 $\mathcal{R}=\{1,\cdots,1\}$.

由于 $\boldsymbol{U}$ 的任意两行的相遇数为常数, 则易验证所构造的 $\mathrm{URBD}(n,\mathcal{A},\mathcal{R})$ 是一个 $\mathrm{URD}(n,\mathcal{A},\lambda,\mathcal{R})$, 其中 $\lambda=\left(\sum_{j=1}^{s}n/q_j-s\right)\Big/(n-1)$.

注意到, Liu 和 Fang(2005) 研究了一类可分解不完全区组设计的均匀性, 并提

出了一种构造这类均匀区组设计的方法. 事实上, 他们构造的均匀区组设计就是这里的 URD, 上述方法就是他们提出的.

归纳一下, 我们有下述定理，它在上述构造方法中扮演着重要角色.

定理 3.5　*在离散偏差准则下, 存在一个均匀设计* $U_n(q_1^{r_1}\times\cdots\times q_l^{r_l})$, *记为* $\boldsymbol{U}$, $\boldsymbol{U}$ *对所有* $1\leqslant i\neq j\leqslant n$ *都有* $\delta_{ij}(\boldsymbol{U})=\delta$ *当且仅当存在一个* $\mathrm{URD}(n,\mathcal{A},\delta,\mathcal{R})$, *其中* $\mathcal{A}=\{n/q_1,\cdots,n/q_l\}$, $\mathcal{R}=\{r_1,\cdots,r_l\}$, *且* $\delta=\left[\sum_{j=1}^{l}r_j(n/q_j-1)\right]\Big/(n-1)$.

通过例 3.12, 说明了 $\mathrm{URD}(n,\mathcal{A},\delta,\mathcal{R})$ 与任意两行的相遇数为常数的均匀设计 $U_n(q_1^{r_1}\cdots q_l^{r_l})$ 之间的等价性. 很容易看出这种构造均匀设计的方法有许多优点:

1) 我们可以找到许多均匀设计而不需要任何计算搜索.

2) 该方法既可构造对称均匀设计也可以构造非对称均匀设计.

3) 该方法可用于构造超饱和均匀设计. 后者将在第 7 章中介绍.

在因子设计中, 如果一列可以通过另一列水平置换得到, 那么这两列称为是*完全别名的* (fully aliased). 对给定的 (n,s,q), 由于不能将完全别名的两列安排给两个不同因子, 故希望所得的均匀设计的任意两列都不是完全别名的. 因此在 $\mathcal{U}(n;q^s)$ 下构造均匀设计的时候, 对于任两列, 需要限制来自 q^2 个水平组合中的每一个水平组合的数目来避免完全别名的列的出现. 在均匀设计的构造中, 将始终坚持这一观点, 特别是均匀设计的组合构造.

在结束这一小节之前, 让我们来看另一个用 URD 构造均匀设计的例子.

例 3.13　*假设有一个* $\mathrm{URD}(12,\{2,3\},1,\{5,3\})$, *其处理集为* $\mathcal{V}=\{1,2,3,4,5,6,7,8,9,10,11,12\}$, *八个平行类如下:*

$$P_1=\{\{1,2\},\{3,4\},\{5,6\},\{7,8\},\{9,10\},\{11,12\}\};$$

$$P_2=\{\{1,3\},\{2,4\},\{5,7\},\{6,8\},\{9,11\},\{10,12\}\};$$

$$P_3=\{\{1,4\},\{2,3\},\{5,8\},\{6,7\},\{9,12\},\{10,11\}\};$$

$$P_4=\{\{1,7\},\{2,8\},\{3,12\},\{4,9\},\{5,10\},\{6,11\}\};$$

$$P_5=\{\{3,5\},\{4,6\},\{1,10\},\{2,11\},\{7,12\},\{8,9\}\};$$

$$P_6=\{\{1,5,9\},\{2,6,10\},\{3,7,11\},\{4,8,12\}\};$$

$$P_7=\{\{1,6,12\},\{2,7,9\},\{3,8,10\},\{4,5,11\}\};$$

$$P_8=\{\{1,8,11\},\{2,5,12\},\{3,6,9\},\{4,7,10\}\}.$$

遵循算法 3.6, 可由该 URD 构造表 3.6 中的 $U_{12}(6^54^3)$. *反过来, 由算法 3.7, 也可用* $U_{12}(6^54^3)$ *来构造相应的* $\mathrm{URD}(12,\{2,3\},1,\{5,3\})$.

表 3.6 一个 $U_{12}(6^5 4^3)$

行	1	2	3	4	5	6	7	8
1	1	1	1	1	3	1	1	1
2	1	2	2	2	4	2	2	2
3	2	1	2	3	1	3	3	3
4	2	2	1	4	2	4	4	4
5	3	3	3	5	1	1	4	2
6	3	4	4	6	2	2	1	3
7	4	3	4	1	5	3	2	4
8	4	4	3	2	6	4	3	1
9	5	5	5	4	6	1	2	3
10	5	6	6	5	3	2	3	4
11	6	5	6	6	4	3	4	1
12	6	6	5	3	5	4	1	2

3.6.2 通过组合设计构造均匀设计

算法 3.6 和定理 3.5 确保我们可以通过 URD 构造均匀设计, 组合设计理论中对后者已有广泛的研究. 本小节将会介绍几种组合设计 (其中大部分都是 URD), 并且利用这些组合设计和算法 3.6 来构造均匀设计. 此外, 将会给出一些例子来说明这些区组设计. 具体地, A—E 介绍对称均匀设计的构造方法, 而 F—H 介绍非对称均匀设计的构造方法. 除了本小节介绍的方法外, 还有其他的一些关于 URD 存在性的结果, 读者可以参见 (Fang et al., 2001, 2004a) 及其相关参考文献.

A. 通过可分解平衡不完全区组设计构造

根据前面介绍的 RBIBD 的定义, 可知 RBIBD(n,k,λ) 实际上是一个特殊的 URD$(n,\mathcal{A},\lambda,\mathcal{R})$, 其中 $\mathcal{A}=\{k\}$, $\mathcal{R}=\{s\}$, 且由 (3.16) 得 $s=\lambda(n-1)/(k-1)$. 因此, 可以用 RBIBD 来构造对称均匀设计 $U_n\left(\binom{n}{k}^s\right)$, 具体细节可参见 (Fang et al., 2002b, 2003a, 2004a). 他们的工作推广了 (Nguyen, 1996) 与 (Liu and Zhang, 2000) 的结果. 现将他们的结果总结成下面的定理.

定理 3.6 假设采用离散偏差作为均匀性度量. 下述 $U_n(q^s)$ 可以通过算法 3.6 构造出来, 其中 n, m, k 都是正整数且它们的取值依具体情况而定.

(a) 若 $n=2m$ 是偶数, 则存在 $U_n(m^{k(n-1)})$, 其中 k 是一个正整数.

(b) 若 $n=6m+3$, 则存在 $U_n((2m+1)^{\frac{n-1}{2}})$.

(c) 若 $n=3m$ 且 $n\neq 6$, 则存在 $U_n(m^{n-1})$.

(d) 若 $n=12m+4$, 则存在 $U_n((3m+1)^{\frac{n-1}{3}})$.

(e) 若 $n=4m$, 则存在 $U_n(m^{n-1})$.

(f) 若 $n=6m$ 且 $n\neq 174, 240$, 则存在 $U_n(m^{n-1})$.

(g) 若 $n = 6m$, 则存在 $U_n(m^{2(n-1)})$.

(h) 若 $n = 20m + 5$ 且 $n \neq 45, 225, 345, 465, 645$, 则存在 $U_n((m+1)^{\frac{n-1}{4}})$.

(i) 若 $n = 5m$ 且 $n \neq 10, 15, 70, 90, 135, 160, 190, 195$, 则存在 $U_n(m^{n-1})$.

表 3.7 和表 3.8 展示了一个 RBIBD 的例子及相应的均匀设计.

表 3.7　一个 RBIBD(10, 2, 1)

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
b_1^j	{1,10}	{2,10}	{4,9}	{3,7}	{2,8}	{5,7}	{5,6}	{1,7}	{1,6}
b_2^j	{8,9}	{5,8}	{3,10}	{4,10}	{6,9}	{2,4}	{3,4}	{2,5}	{2,7}
b_3^j	{4,5}	{3,6}	{7,8}	{1,2}	{5,10}	{1,9}	{1,8}	{4,6}	{4,8}
b_4^j	{6,7}	{7,9}	{2,6}	{5,9}	{1,3}	{3,8}	{7,10}	{3,9}	{3,5}
b_5^j	{2,3}	{1,4}	{1,5}	{6,8}	{4,7}	{6,10}	{2,9}	{8,10}	{9,10}

注: b_i^j, 第 j 个平行类中的第 i 个区组.

表 3.8　由表 3.7 得到的一个 $U_{10}(5^9)$

行	1	2	3	4	5	6	7	8	9
1	1	5	5	3	4	3	3	1	1
2	5	1	4	3	1	2	5	2	2
3	5	3	2	1	4	4	2	4	4
4	3	5	1	2	5	2	2	3	3
5	3	2	5	4	3	1	1	2	4
6	4	3	4	5	2	5	1	3	1
7	4	4	3	1	5	1	4	1	2
8	2	2	3	5	1	4	3	5	3
9	2	4	1	4	2	3	5	4	5
10	1	1	2	2	3	5	4	5	5

B. 通过 Room 方构造

Room 方 (room square) 是组合设计理论中一个很重要的概念. 有关其全面的介绍可参考 (Colbourn and Dinita, 1996). Fang 等 (2002a) 将 Room 方用于对称超饱和设计的构造, 所得设计实际上是均匀设计. 下面给出 Room 方的定义.

定义 3.8　令 $\mathcal{V}$ 是一个有 n 个元素 (处理) 的集合. 一个 (在处理集 $\mathcal{V}$ 上的) $n-1$ 阶 Room 方是一个 $(n-1)\times(n-1)$ 的阵列 $\mathcal{F}$, 其满足如下性质:

1) $\mathcal{F}$ 中的每一个位置要么是空的, 要么包含集合 $\mathcal{V}$ 中的无序处理对.

2) $\mathcal{V}$ 中的每个元素在 $\mathcal{F}$ 中的每行和每列中均恰好出现一次.

3) 集合 $\mathcal{V}$ 中的无序处理对恰好在 $\mathcal{F}$ 中出现一次.

表 3.9 给出了一个 7 阶 Room 方. 通过该表和定义 3.8, 可以很容易地观察到, 如果将无序处理对作为区组, 那么 $n-1$ 阶 Room 方的每一行构成的区组都是一个

平行类, 而这 $n-1$ 个平行类可以组成一个 RBIBD$(n,2,1)$. 同样地, 用该 Room 方的每一列作为平行类又可以得到另一个 RBIBD$(n,2,1)$. 然后将这两个 RBIBD 合在一起, 可以得到一个 RBIBD$(n,2,2)$. 因此, 可以得到如下定理.

表 3.9 一个 7 阶 Room 方

81			26		57	34
45	82			37		61
72	56	83			41	
	13	67	84			52
63		24	71	85		
	74		35	12	86	
		15		46	23	87

定理 3.7 给定一个 $n-1$ 阶 Room 方, 就可以得到两个 RBIBD$(n,2,1)$ 和一个 RBIBD$(n,2,2)$, 从而可以通过算法 3.6 构造出两个 $U_n\left(\left(\frac{n}{2}\right)^{n-1}\right)$ 和一个 $U_n\left(\left(\frac{n}{2}\right)^{2n-2}\right)$, 同时, 这些均匀设计中不存在完全别名的列.

表 3.10 列出了通过表 3.9 中的 Room 方构造出的两个均匀设计 $U_8(4^7)$ 和一个均匀设计 $U_8(4^{14})$(分别列在 $1,\cdots,7$ 列, $8,\cdots,14$ 列和 $1,\cdots,14$ 列). 关于该方法的更多结果可参见 (Fang et al., 2002a).

表 3.10 由表 3.9 得到的均匀设计

行	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	4	4	1	3	3	1	1	3	4	3	3	2	2
2	2	2	1	4	2	3	3	3	1	3	1	3	4	3
3	4	3	3	1	1	2	3	4	3	1	4	1	4	1
4	4	1	4	3	2	1	2	2	4	3	2	4	2	1
5	3	1	2	4	4	2	1	2	2	4	4	2	1	3
6	2	4	2	2	1	4	2	4	2	2	1	4	3	2
7	3	3	1	2	3	1	4	3	4	2	3	1	1	4
8	1	2	3	3	4	4	4	1	1	1	2	2	3	4

C. 通过可分解填充设计构造

可分解填充设计在组合设计理论中已有广泛的研究. Fang 等 (2004b) 建立了可分解填充设计与超饱和设计之间的紧密联系, 并提出了用这种设计构造超饱和设计的方法. 实际上, 他们这种方法构造的超饱和设计就是离散偏差意义下的均匀设计. 下面介绍填充设计的相关概念.

定义 3.9 设 $n \geqslant p \geqslant t$. 一个 $t-(n,p,1)$ 填充设计 (packing design) 是一

对 $(\mathcal{V},\mathcal{B})$, 其中 $\mathcal{V}$ 是一个 n 元 (处理) 集合, $\mathcal{B}$ 由 $\mathcal{V}$ 的 p 元子集 (区组) 构成, 满足 $\mathcal{V}$ 的任一 t 元子集至多出现在 $\mathcal{B}$ 的一个区组中. 称可分解 t-$(n,p,1)$ 填充设计中区组能够达到的最大个数为填充数 (packing number), 记为 $N(n,p,t)$. 如果 $|\mathcal{B}|=N(n,p,t)$, 则称 t-$(n,p,1)$ 填充设计 $(\mathcal{V},\mathcal{B})$ 是最优的.

有关填充设计的更多讨论, 请参见文献 (Stinson, 1996). 特别地, Stinson(1996) 给出了如下结论.

定理 3.8 ((Stinson, 1996) 中的定理 33.5)　填充数的上界如下:

$$N(n,p,t) \leqslant \left\lfloor \frac{n}{p} \left\lfloor \frac{n-1}{p-1} \cdots \left\lfloor \frac{n-t+1}{p-t+1} \right\rfloor \right\rfloor \right\rfloor.$$

由定义 3.9 容易看出, 给定一个可分解的填充设计 (resolvable packing design), 可以通过算法 3.6 构造出一个 U-型设计. 表 3.11 给出了一个可分解的 2-$(6,2,1)$ 填充设计 $(\mathcal{V},\mathcal{B})$, 其中 $\mathcal{V}=\{1,2,3,4,5,6\}$, $\mathcal{B}$ 被分成 5 个平行类, 每个平行类由 3 个大小为 2 的不相交的区组构成, 并且每一个无序的二元对恰好出现在 $\mathcal{B}$ 的一个区组中. 注意, 该设计是一个最优的 2-$(6,2,1)$ 填充设计, 因为 $\mathcal{B}$ 包含了所有大小为 2 的不同区组, 再多加一个区组, 如 $\{k,l\}$, 到 $\mathcal{B}$ 中会使得二元对 $\{k,l\}$ 出现在 $\mathcal{B}$ 的两个区组中. 表 3.12 给出了由该最优填充设计构造的 U-型设计. 实际上它也是一个均匀设计 $U_6(3^5)$.

表 3.11　一个可分解的 2-(6, 2, 1) 填充设计

$\mathcal{V}$	$\{1,2,3,4,5,6\}$				
$\mathcal{B}$	P_1	P_2	P_3	P_4	P_5
b_1^j	{1,2}	{1,3}	{1,4}	{1,5}	{1,6}
b_2^j	{3,4}	{2,5}	{2,6}	{2,4}	{2,3}
b_3^j	{5,6}	{4,6}	{3,5}	{3,6}	{4,5}

注: b_i^j, 第 j 个平行类中的第 i 个区组.

表 3.12　由表 3.11 得到的 $U_6(3^5)$

行	1	2	3	4	5
1	1	1	1	1	1
2	1	2	2	2	2
3	2	1	3	3	2
4	2	3	1	2	3
5	3	2	3	1	3
6	3	3	2	3	1

然而一般来说, 不能保证所得到的 U-型设计的任两行之间的相遇数是常数, 因为不能确定每个处理对是否都出现在填充设计的相同个数的区组中. 因此, 自然的

想到要考察由可分解的填充设计构造出的 U-型设计的均匀性及其他性质. 为了回答这个问题, 需要先给出最大频数的定义.

定义 3.10 令 $\mathcal{U}(n;q^s;r)$ 是 $\mathcal{U}(n;q^s)$ 的一个子集, 要求 $\mathcal{U}(n;q^s;r)$ 中的任一设计满足: 任两列的任一水平组合最多出现 r 次. 称 r 为设计的最大频数 (largest frequency).

根据定义 3.9 和定义 3.10, Fang 等 (2004b) 得到了 $\mathcal{U}(n;q^s;r)$ 中的设计与可分解的填充设计之间的联系.

定理 3.9 存在一个属于 $\mathcal{U}(n;q^s;r)$ 的设计等价于存在一个包含 s 个平行类的可分解的 $(r+1)$-$(n,p,1)$ 填充设计, 其中 $2 \leqslant p = n/q \leqslant q$ 且 $1 \leqslant r < p$.

基于该联系和定理 3.8 给出的填充数的上界, Fang 等 (2004b) 给出了 $\mathcal{U}(n;q^s;r)$ 中设计列数 s 的上界.

定理 3.10 对给定的满足 $2 \leqslant p = n/q \leqslant q$ 和 $1 \leqslant r < p$ 的 (n,q,r), s 的上界由下式给出

$$s \leqslant \left\lfloor \frac{n-1}{p-1} \left\lfloor \frac{n-2}{p-2} \cdots \left\lfloor \frac{n-r}{p-r} \right\rfloor \right\rfloor \right\rfloor .$$

注意到对 $\mathcal{U}(n;q^s;1)$ 中的任一设计, 任两行之间的相遇数最多为1. 从而根据定理 2.23、定理 3.9 和定理 3.10, 有如下定理.

定理 3.11 对于 $2 \leqslant p = n/q \leqslant q$ 且 $s \leqslant \lfloor (n-1)/(p-1) \rfloor$, 每一个设计 $\boldsymbol{U} \in \mathcal{U}(n;q^s;1)$ 都是一个 $U_n(q^s)$, 并且可以由可分解的 2-$(n,p,1)$ 填充设计构造得到.

该定理告诉我们 $\mathcal{U}(n;q^s;1)$ 中的任一设计都是均匀的, 无论 s 是否达到了定理 3.10 中的上界. 若 s 达到了上界, 则相应的可分解的填充设计是最优的.

D. 通过 Kirkman 三元系大集构造

Kirkman 三元系大集 (large set of Kirkman triple systems) 在组合设计理论中是一个重要的研究对象 (参见 (Stinson, 1991)), 它可以看作一种可分解的最优填充设计, 也可以用来构造 $\mathcal{U}(n;q^s;2)$ 中的均匀设计.

定义 3.11 一个 n 阶的 Steiner 三元系 (Steiner triple system, STS), 记为 STS(n), 是一对 $(\mathcal{V},\mathcal{B})$, 其中 $\mathcal{V}$ 是一个 n 元 (处理) 集合, $\mathcal{B}$ 由 $\mathcal{V}$ 的三元子集 (称为三元组或区组) 构成, 满足 $\mathcal{V}$ 中任一无序的二元对恰好出现在 $\mathcal{B}$ 的一个三元组中. 若 $\mathcal{B}$ 是可分解的, 则称 STS(n) 为一个 Kirkman 三元系 (Kirkman triple system, KTS), 记为 KTS(n). 同一集合 $\mathcal{V}$ 上 $(n-2)$ 个两两不相交的 KTS(n) 构成 KTS(n) 的一个大集 (large set), 记为 LKTS(n).

注意的是由该定义, 一个 KTS(n) 实际上是一个可分解的最优 2-$(n,3,1)$ 填充设计. 在这样一个设计中, 恰好有 $n(n-1)/6$ 个三元组, 它们包含了所有的 $\begin{pmatrix} n \\ 2 \end{pmatrix}$ 个

无序的二元对, 而所有的三元组被分成了 $(n-1)/2$ 个平行类. 因为一个 LKTS(n) 包含了 $\mathcal{V}$ 中所有的 $\begin{pmatrix} n \\ 3 \end{pmatrix} = (n-2)[n(n-1)/6]$ 个不同的三元组, 所以它是一个可分解的最优 3-$(n,3,1)$ 填充设计. 同样还可以看出一个 KTS(n) 是一个 RBIBD(n, 3, 1), 并且一个 LKTS(n) 是由 $(n-2)$ 个 RBIBD($n,3,1$) 构成的 RBIBD($n,3,n-2$). 在这种情况下, $q=n/3$ 是一个正整数. 因此, 由上述讨论可以得到如下定理.

定理 3.12 给定一个 LKTS(n) ($n \geqslant 9$), 令 $q=n/3$ 且 $m=(n-1)/2$, 则

(1) 可以通过算法 3.6 构造 $\mathcal{U}(n;q^m;1)$ 中的 $(n-2)$ 个 $U_n(q^m)$ 和 $\mathcal{U}(n;q^{(n-2)m};2)$ 中的一个 $U_n(q^{(n-2)m})$.

(2) $U_n(q^{(n-2)m})$ 的列数和任一个 $U_n(q^m)$ 的列数都达到了定理 3.10 中给出的上界.

(3) 对 $1<l<n-2$, 可以通过任意 l 个 $U_n(q^m)$ 构造 $\mathcal{U}(n;q^{lm};2)$ 中的一个 $U_n(q^{lm})$, 并且对于给定的 $(n,q,s)=(n,n/3,lm)$, 2 是 $\mathcal{U}(n;q^s;r)$ 的参数 r 的最小值.

表 3.13 和表 3.14 给出了一个 LKTS(9) 的例子及对应的均匀设计. 表 3.14 给出的 $U_9(3^{28})$ 共有 7 部分, 每一部分有 4 列. 这一设计有下面这些性质:

(a) 每一部分是一个 $L_9(3^4)$, 即 $r=1$;

(b) 对任一 $l=2,\cdots,7$, 由前 $4l$ 列构成的设计是一个 $U_9(3^{4l})$ 且有最大频数 $r=2$;

(c) 对给定的 $(n,q,r)=(9,3,2)$, s 的最大值为 28, 且被 $U_9(3^{28})$ 达到;

(d) 对给定的 $(n,q,s)=(9,3,4l), l=2,\cdots,7$, r 的最小值为 2, 且被 $U_9(3^{4l})$ 达到.

表 3.13 一个 LKTS(9)

$\mathcal{V}$	{1, 2, 3, 4, 5, 6, 7, 8, 9}			
	P_1^i	P_2^i	P_3^i	P_4^i
$\mathcal{B}_1$	{{124}{356}{789}}	{{257}{468}{139}}	{{347}{158}{269}}	{{167}{238}{459}}
$\mathcal{B}_2$	{{235}{467}{189}}	{{361}{578}{249}}	{{451}{268}{379}}	{{271}{348}{569}}
$\mathcal{B}_3$	{{346}{571}{289}}	{{472}{618}{359}}	{{562}{378}{419}}	{{312}{458}{679}}
$\mathcal{B}_4$	{{457}{612}{389}}	{{513}{728}{469}}	{{673}{418}{529}}	{{423}{568}{719}}
$\mathcal{B}_5$	{{561}{723}{489}}	{{624}{138}{579}}	{{714}{528}{639}}	{{534}{678}{129}}
$\mathcal{B}_6$	{{672}{134}{589}}	{{735}{248}{619}}	{{125}{638}{749}}	{{645}{718}{239}}
$\mathcal{B}_7$	{{713}{245}{689}}	{{146}{358}{729}}	{{236}{748}{159}}	{{756}{128}{349}}

注: {124}表示区组 {1, 2, 4}, 以此类推.

表 3.14 由表 3.13 得到的 $U_9(3^{4s})(1 \leqslant s \leqslant 7)$

行	1				2				3				4				5				6				7			
1	1	3	2	1	3	1	1	1	2	2	3	1	2	1	2	3	1	2	1	3	2	3	1	2	1	1	3	2
2	1	1	3	2	1	3	2	1	3	1	1	1	2	2	3	1	2	1	2	3	1	2	1	3	2	3	1	2
3	2	3	1	2	1	1	3	2	1	3	2	1	3	1	1	1	2	2	3	1	2	1	2	3	1	2	1	3
4	1	2	1	3	2	3	1	2	1	1	3	2	1	3	2	1	3	1	1	1	2	2	3	1	2	1	2	3
5	2	1	2	3	1	2	1	3	2	3	1	2	1	1	3	2	1	3	2	1	3	1	1	1	2	2	3	1
6	2	2	3	1	2	1	2	3	1	2	1	3	2	3	1	2	1	1	3	2	1	3	2	1	3	1	1	1
7	3	1	1	1	2	2	3	1	2	1	2	3	1	2	1	3	2	3	1	2	1	1	3	2	1	3	2	1
8	3	2	2	2	3	2	2	2	3	2	2	2	3	2	2	2	3	2	2	2	3	2	2	2	3	2	2	2
9	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

注意到这些设计都是均匀设计. 这是一个很神奇的结果! Fang 等 (2004b) 进一步指出, 在其他某些准则下, 例如, 将在 7.2 节介绍的 $E(f_{\text{NOD}})$, 这些设计比 Yamada 等 (1999) 与 Fang 等 (2000a) 通过计算机搜索得到的设计表现更好. 关于由可分解的填充设计和 LKTS 得到的设计得更详细结果可参见 (Fang et al., 2004b).

E. 通过超单可分解的 t-设计构造

超单可分解的 t-设计可用来构造 $\mathcal{U}(n;q^s;r)$ 中的均匀设计. 先介绍有关可分解的 t-设计的一些概念.

定义 3.12 一个 t-设计 (t-design), 记为 $S_\lambda(t,k,n)$, 是一个有序对 $(\mathcal{V},\mathcal{B})$, 其中 $\mathcal{V}$ 是一个 n 元 (处理) 集合, $\mathcal{B}$ 由 $\mathcal{V}$ 的 k 元子集 (称为区组) 构成, 满足 $\mathcal{V}$ 的每一个 t 元子集恰好在 $\mathcal{B}$ 的 λ 个区组中出现. 当 $t=2$ 时, 一个 $S_\lambda(t,k,n)$ 就是一个平衡不完全区组设计 (balanced incomplete block design). 若一个 $S_\lambda(t,k,n)$ 没有重复区组, 则称之为单纯的 (simple). 若一个单纯 $S_\lambda(t,k,n)$ 的任两个区组间的相同点都不多于两个, 则称之为超单的 (super-simple). 若一个 $S_\lambda(t,k,n)$ 是可分解的, 则记为 $RS_\lambda(t,k,n)$.

读者可以参考 (Beth et al., 1999) 来了解关于超单 $S_\lambda(t,k,n)$ 的更多讨论. 由该定义, 可以知道, $S_1(2,k,n)$, $S_1(3,4,n)$ 和单纯的 $S_\lambda(2,3,n)$ 都是超单的. 而且, 基于 3.6.1 小节的讨论, Fang 等 (2004c) 建立了设计间的如下关系.

定理 3.13 存在一个属于 $\mathcal{U}(n;q^s)$ 且任意 t 个不同行之间相遇数为 λ 的设计等价于存在一个有 s 个平行类的 $RS_\lambda(t,n/q,n)$.

进一步地, 如果用一个超单 $RS_\lambda(2,k,n)$ 来构造一个 $U(n;q^s)$ 设计, 则容易看出所得到的设计 $U(n;q^s)$ 的最大频数为 2, 也就是说, 该设计属于 $\mathcal{U}(n;q^s;2)$. Fang 等 (2004c) 给出了如下结论.

定理 3.14 (1) 存在一个属于 $\mathcal{U}(n;q^s;2)$ 的设计, 等价于存在一个有 s 个平行类的超单 $RS_\lambda(2,\,n/q,n)$, 其中 $\lambda=s(n-q)/(q(n-1))$.

(2) 给定一个超单 $RS_\lambda(2,k,n)$, 通过算法 3.6 构造的属于 $\mathcal{U}(n;q^s;2)$ 的设计是一个均匀设计且没有完全别名的列.

为了说明该结论, 例 3.14 给出了一个超单 $RS_2(2,3,12)$ 及由它构造出来的均匀设计. 更多结果, 例如, $RS_2(2,3,9)$, $RS_1(3,4,16)$ 及相应的 $U_9(3^8;2)$,$U_{16}(4^{35};2)$ 可以参见 (Fang et al., 2004c).

例 3.14　当 $n=12$ 时, $\mathcal{V}=\{1,2,3,4,5,6,7,8,9,10,11,12\}$ 及一个给定的平行类 P_1, 其中该平行类包含 4 个区组 $b_1^1=\{1,2,4\}$,$b_2^1=\{3,6,8\}$,$b_3^1=\{5,9,10\}$ 和 $b_4^1=\{7,11,12\}$. 将 $(1,2,3,4,5,6,7,8,9,10,11)(12)$ 生成的置换群应用到该平行类上, 可以得到各包含 4 个区组的 11 个平行类. 例如, 根据 $(1,2,3,4,5,6,7,8,9,10,11)(12)$ 生成的置换群, 对 P_1 中的每个元素加 1 得到平行类 $P_2=\{\{2,3,5\},\{4,7,9\},\{6,10,11\},\{8,1,12\}\}$. 此时, $\mathcal{V}$ 中的每个无序对恰好出现在两个区组中, 同时 $\mathcal{V}$ 的任一个三元组至多出现在一个平行类中. 这样就得到了一个超单 $RS_2(2,3,12)$, 见表 3.15. 由该超单 $RS_2(2,3,12)$, 可用算法 3.6 构造出一个均匀设计 $U_{12}(4^{11};2)$, 同样列于表 3.15 中. 容易将该例子推广到一般情况.

表 3.15　一个均匀设计 $U_{12}(4^{11};2)$ 及相应的单纯 $RS_2(2,3,12)$

行	1	2	3	4	5	6	7	8	9	10	11
1	1	4	3	3	2	4	2	3	1	2	1
2	1	1	4	3	3	2	4	2	3	1	2
3	2	1	1	4	3	3	2	4	2	3	1
4	1	2	1	1	4	3	3	2	4	2	3
5	3	1	2	1	1	4	3	3	2	4	2
6	2	3	1	2	1	1	4	3	3	2	4
7	4	2	3	1	2	1	1	4	3	3	2
8	2	4	2	3	1	2	1	1	4	3	3
9	3	2	4	2	3	1	2	1	1	4	3
10	3	3	2	4	2	3	1	2	1	1	4
11	4	3	3	2	4	2	3	1	2	1	1
12	4	4	4	4	4	4	4	4	4	4	4
$\mathcal{B}$	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}
b_1^j	$\{124\}$	$\{235\}$	$\{346\}$	$\{457\}$	$\{568\}$	$\{679\}$	$\{78t_0\}$	$\{89t_1\}$	$\{9t_01\}$	$\{t_0t_12\}$	$\{t_113\}$
b_2^j	$\{368\}$	$\{479\}$	$\{58t_0\}$	$\{69t_1\}$	$\{7t_01\}$	$\{8t_12\}$	$\{913\}$	$\{t_024\}$	$\{t_135\}$	$\{146\}$	$\{257\}$
b_3^j	$\{59t_0\}$	$\{6t_0t_1\}$	$\{7t_11\}$	$\{812\}$	$\{923\}$	$\{t_034\}$	$\{t_145\}$	$\{156\}$	$\{267\}$	$\{378\}$	$\{489\}$
b_4^j	$\{7t_1t_2\}$	$\{81t_2\}$	$\{92t_2\}$	$\{t_03t_2\}$	$\{t_14t_2\}$	$\{15t_2\}$	$\{26t_2\}$	$\{37t_2\}$	$\{48t_2\}$	$\{59t_2\}$	$\{6t_0t_2\}$

注: b_i^j, 第 j 个平行类中的第 i 个区组; t_i:$10+i$, 其中 $i=0,1,2$.

F. 通过可分解的可分组设计构造

Fang 等 (2001) 用可分解的可分组设计(resolvable group divisible design, RGDD) 来构造对称和非对称 (混水平) 均匀设计.

定义 3.13 令 k 和 g 为正整数, n 为 g 的整数倍. 指数 (index) 为 1, 组型 (type) 为 $g^{n/g}$ 的 n 阶(order) 可分组设计 (group divisible design, GDD), 记为 k-GDD($g^{n/g}$), 是一个三元组 $(\mathcal{V},\mathcal{G},\mathcal{B})$, 其中:

(1) $\mathcal{V}$ 是 n 个处理的一个集合;

(2) $\mathcal{G}$ 是 $\mathcal{V}$ 的一个划分, 划分后每组大小为 g;

(3) $\mathcal{B}$ 是一族 $\mathcal{V}$ 中大小为 k 的区组;

(4) 每个处理对恰好出现在一个区组或组中, 但不同时出现在区组和组中.

若 GDD 中的区组可以被分成若干平行类, 则称该 GDD 是一个可分解的 GDD, 记为 k-RGDD($g^{n/g}$). 显然当 $g=1$ 时, 一个 k-RGDD 其实就是一个 RBIBD$(n,k,1)$.

例 3.15 由例 3.11 中的 URD$(6,\{3,2\},1,\{1,3\})$, 取 $\mathcal{V}=\{1,2,3,4,5,6\},\mathcal{G}=P_1$ 及 $\mathcal{B}=\bigcup_{i=2}^{4}P_i$, 可以得到一个 2-RGDD($3^2$).

假设 $(\mathcal{V},\mathcal{G},\mathcal{B})$ 是一个 k-RGDD($g^{n/g}$). 令 $\mathcal{B}^*=\mathcal{G}\bigcup\mathcal{B}$, 则可将 $\mathcal{B}^*$ 看作一个区组设计, 其中 $\mathcal{G}$ 中的组构成区组大小为 g 的一个均匀平行类. 由 GDD 的定义可以看出, $\mathcal{V}$ 的任一元素对都恰好出现在 $\mathcal{B}^*$ 的一个区组中. 因此, 由初始 RGDD 的可分解性和均匀性可知, 区组设计 $\mathcal{B}^*$ 是一个 URD 并且可以用来构造均匀设计. Fang 等 (2001) 给出了下述定理和许多新的均匀/超饱和设计.

定理 3.15 若存在一个 k-RGDD($g^{n/g}$), 则存在一个 URD$(n,\{g,k\},1,\{1,r_k\})$, 其中 $r_k=(n-g)/(k-1)$, 从而存在一个 $U_n\left(\left(\frac{n}{g}\right)^1\left(\frac{n}{k}\right)^{\frac{n-g}{k-1}}\right)$.

G. 通过拉丁方构造

拉丁方矩阵在试验设计中长期占据重要地位. 这里将利用拉丁方来构造 k-RGDD(g^k) 设计, 同样也是均匀设计. 首先回顾一下拉丁方的定义.

定义 3.14 称一个由 g 个符号作为其元素构成的 $g\times g$ 矩阵为 g 阶拉丁方 (Latin square), 记为 $L=(L_{ij})$, 若它满足每个符号分别在每行和每列中出现且仅出现一次. 称该性质为拉丁性质 (Latin property). 若两个拉丁方叠加后生成 g^2 个不同的有序对, 则称这两个拉丁方是正交的 (orthogonal). 对于一个拉丁方的集合, 如果其中的任意一对拉丁方都是正交的, 则称该拉丁方集合是两两正交拉丁方 (pairwise orthogonal Latin squares) 集合.

用 $N(g)$ 表示 g 阶两两正交拉丁方的最大个数. 关于 $N(g)$ 的已有结论, 读者可以参考 (Colbourn and Dinita, 1996). Fang 等 (2001) 给出了以下结论.

定理 3.16　若 $N(g) \geqslant k-1$, 则存在一个 k-RGDD(g^k), 进而可以得到一个试验次数为 $n=kg$、因子数为 $g+1$ 的设计 $U_{kg}(k^1 g^g)$, 其中 1 个因子是 k 水平, 其余 g 个因子是 g 水平的.

来看一个通过两两正交拉丁方构造 RGDD 和均匀设计的例子.

例 3.16　表 3.16 给出了两个正交的 4 阶拉丁方, 我们用它们可以构造一个 3-RGDD(4^3) $(\mathcal{V},\mathcal{G},\mathcal{B})$ 如下, 其中 $\mathcal{V}=\{1,2,3,4,5,6,7,8,9,10,11,12\}$, $\mathcal{G}=\{\{1,2,3,4\},\{5,6,7,8\},\{9,10,11,12\}\}$, $\mathcal{B}=\bigcup_{i=1}^{4} P_i$ 且

$$P_1=\{\{1,8,9\},\{2,6,11\},\{3,7,12\},\{4,5,10\}\};$$
$$P_2=\{\{1,5,11\},\{2,7,9\},\{3,6,10\},\{4,8,12\}\};$$
$$P_3=\{\{1,6,12\},\{2,8,10\},\{3,5,9\},\{4,7,11\}\};$$
$$P_4=\{\{1,7,10\},\{2,5,12\},\{3,8,11\},\{4,6,9\}\}.$$

由该 RGDD, 可以构造出表 3.17 中的 $U_{12}(3^1 4^4)$.

表 3.16　两个正交的 4 阶拉丁方

设计 1				设计 2			
8	6	7	5	9	11	12	10
5	7	6	8	11	9	10	12
6	8	5	7	12	10	9	11
7	5	8	6	10	12	11	9

表 3.17　一个 $U_{12}(3^1 4^4)$

行	1	2	3	4	5
1	1	1	1	1	1
2	1	2	2	2	2
3	1	3	3	3	3
4	1	4	4	4	4
5	2	4	1	3	2
6	2	2	3	1	4
7	2	3	2	4	1
8	2	1	4	2	3
9	3	1	2	3	4
10	3	4	3	2	1
11	3	2	1	4	3
12	3	3	4	1	2

H. 通过可分解的部分成对平衡设计构造

Fang 等 (2006c) 提出了利用可分解的部分成对平衡设计 (resolvable partially pairwise balanced design) 构造均匀设计, 所构造的均匀设计的任两行之间的 Hamming 距离最多相差 1. 先让我们回顾一下相关术语.

定义 3.15 令 $\mathcal{V}$ 是一个有 n 个处理的集合, $\mathcal{B}$ 是一族 $\mathcal{V}$ 中的区组, 区组大小来自集合 $\mathcal{A}$, 满足任一处理对恰好出现在 λ_1 或 λ_2 个区组中. 称该设计为区组大小集合为 $\mathcal{A}$ 的 n 阶部分成对平衡设计 (partially pairwise balanced design), 记为 $(n, \mathcal{A}; \lambda_1, \lambda_2)$-PPBD. 同时记一个有互不相同均匀平行类的可分解 PPBD 为 $\tilde{\text{R}}$PPBD (参见定义 3.6).

当 $\lambda_1 = \lambda_2$ 时, 一个 $\tilde{\text{R}}$PPBD 是有着互不相同平行类的 URD. 需要强调的是, 对给定的参数 n, $\mathcal{A} = \{k_1, \cdots, k_l\}$, λ_1 和 λ_2, 一个 $(n, \mathcal{A}; \lambda_1, \lambda_2)$-$\tilde{\text{R}}$PPBD 的平行类个数并不是常数, 而是取决于它的结构. 基于此, 当一个 $\tilde{\text{R}}$PPBD 含 r_i 个区组大小为 k_i $(i = 1, \cdots, l)$ 的平行类时, 则称它有平行类的型 (class type) $k_1^{r_1} \cdots k_l^{r_l}$.

例 3.17 假设 $\mathcal{V} = \{1, \cdots, 8\}$. 令 $\lambda_1 = 2$, $\lambda_2 = 1$ 且 $\mathcal{A} = \{4, 2\}$. 则下面的区组构成了一个平行类的型为 $4^4 2^1$ 的 $(8, \{4, 2\}; 2, 1)$-$\tilde{\text{R}}$PPBD.

$$\begin{array}{ll}
\{1,3,5,7\}, & \{2,4,6,8\}, \\
\{1,3,6,8\}, & \{2,4,5,7\}, \\
\{1,4,5,8\}, & \{2,3,6,7\}, \\
\{1,4,6,7\}, & \{2,3,5,8\},
\end{array}$$
$$\{1,2\}, \quad \{3,4\}, \quad \{5,6\}, \quad \{7,8\}.$$

依据算法 3.6, 给定一个 $\tilde{\text{R}}$PPBD, 可以生成一个 U-型设计, 并且容易看出所得设计任两行之间的相遇数只取 λ_1和 λ_2 两个值. 因此, 基于定理 2.23, 可以得到如下定理.

定理 3.17 给定一个平行类的型为 $k_1^{r_1} \cdots k_l^{r_l}$ 且满足 $|\lambda_1 - \lambda_2| \leqslant 1$ 的 $(n, \{k_1, \cdots, k_l\}; \lambda_1, \lambda_2)$-$\tilde{\text{R}}$PPBD, 则可通过算法 3.6 构造一个均匀设计 $U_n\left(\left(\dfrac{n}{k_1}\right)^{r_1} \cdots \left(\dfrac{n}{k_l}\right)^{r_l}\right)$, 且该设计中没有完全别名的列.

例 3.18 将算法 3.6 用于例 3.17 给出的 $\tilde{\text{R}}$PPBD, 可以得到下面列出的均匀设计 $U_8(2^4 4^1)$. 有趣的是, 该设计实际上是一个强度为 2 的非对称正交表.

行	1	2	3	4	5
1	1	1	1	1	1
2	2	2	2	2	1
3	1	1	2	2	2
4	2	2	1	1	2
5	1	2	1	2	3
6	2	1	2	1	3
7	1	2	2	1	4
8	2	1	1	2	4

3.6.3 通过饱和正交表构造均匀设计

饱和正交表具有一些有趣的特性, 可用于超饱和与均匀设计的构造. Lin(1993) 提出了一种用 Hadamard 阵的一半 (简记为 HFHM) 来构造大小为 $(n, s) = (2t, 4t-2)$ 的二水平超饱和设计的方法. 之后, Cheng(1997) 给出了 HFHM 的理论依据. 下述定理在这种构造方法中起着重要作用, 并且可以在 (Fang et al., 2003c) 中找到其在更一般情形的推广.

定理 3.18 假设 $\boldsymbol{U}$ 是一个饱和正交表 $L_n(q^s)$, 其中

情形 (i) q 是一个素数幂, $n = q^t$, $s = (n-1)/(q-1)$ 且 $t \geqslant 2$, 或者

情形 (ii) $q = 2$, $n = 4t$, $s = 4t-1$ 且 $t \geqslant 1$,

则该设计任意两行间的 Hamming 距离为常数, 其中在情形 (i) 为 q^{t-1}, 在情形 (ii) 为 $2t$. 也就是说, 在离散偏差准则下, 该设计是均匀设计.

Hadamard 阵是参数满足定理 3.18 中情形 (ii) 的饱和正交设计. Fang 等 (2003c) 将 HFHM 方法推广得到饱和正交表拆分 (fractions of saturated orthogonal arrays, FSOA) 方法, 用于由饱和正交表来构造非对称均匀设计. 下面给出 FSOA 方法.

算法 3.8 (FSOA 方法)

步骤 1 令 $\boldsymbol{U}$ 是一个强度为 2 的饱和正交表. 选定 $\boldsymbol{U}$ 的任意一列, 不妨假定为第 k 列 $(\boldsymbol{k})$, 将 $\boldsymbol{U}$ 的 n 行分成 q 组, 使得第 i 组包含所有在列 $(\boldsymbol{k})$ 中取水平 i 的 $n/q = q^{t-1}$ 行. 这里称列 $(\boldsymbol{k})$ 为分支列 (branching column).

步骤 2 给定 $p(2 \leqslant p < q)$, 任选 p 组生成一个非对称超饱和设计 $U(pq^{t-1}; p^1\, q^{s-1})$, 其中的分支列安排一个 p 水平因子, 其他列安排 $(s-1)$ 个 q 水平因子.

由定理 3.18 可知, 该方法生成的超饱和设计是均匀设计. 现在让我们以表 3.18 中所给出的饱和正交表 $L_{16}(4^5)$ 为例作说明. 如果选列 $(\mathbf{1})$ 作为分支列, 那么 $n = 16$ 行可分为 $q = 4$ 组, 即行 1– 4, 5–8, 9–12 及 13–16. 任意三组可以构成一个非对称超饱和设计 $U_{12}(3^1 4^4)$, 包括 1 个三水平因子和 4 个四水平因子, 例如, 行 1–8 及 13–16. 同样地, 任意两个组可以构成一个非对称超饱和设计 $U_8(2^1 4^4)$, 包括 1 个二水

平因子和 4 个四水平因子, 例如, 行 5–12. 将按上述方法生成的两个设计 $U_{12}(3^1 4^4)$ 及 $U_8(2^1 4^4)$ 列在表 3.18 中.

表 3.18 由 $L_{16}(4^5)$ 生成的均匀设计(以 1 为分支列)

$U_{12}(3^1 4^4)$	$U_8(2^1 4^4)$	行	**1**	**2**	**3**	**4**	**5**
1		1	1	1	1	1	1
2		2	1	2	2	2	2
3		3	1	3	3	3	3
4		4	1	4	4	4	4
5	1	5	2	1	2	3	4
6	2	6	2	2	1	4	3
7	3	7	2	3	4	1	2
8	4	8	2	4	3	2	1
	5	9	3	1	3	4	2
	6	10	3	2	4	3	1
	7	11	3	3	1	2	4
	8	12	3	4	2	1	3
9		13	4	1	4	2	3
10		14	4	2	3	1	4
11		15	4	3	2	4	1
12		16	4	4	1	3	2

关于所构造设计的特性, Fang 等 (2003c) 得到了如下结果.

定理 3.19 *无论选择哪一列作为分支列, 以及选择哪些组构成设计, 由上述 FSOA 方法得到的非对称超饱和设计都是均匀设计, 即 $U_{pq^{t-1}}(p^1 q^{s-1})$, 且 p $(2 \leqslant p < q)$ 水平因子与任一 q 水平因子是正交的, 此外, 所得到的这些均匀设计没有完全别名的列.*

不少设计书籍中提供了强度为 2 的饱和正交表, 例如, (Hedayat et al., 1999). 本书中引用的例子可在网址 http://neilsloane.com/oadir/ 中找到. 因此, 基于 FSOA 方法, 由参数满足定理 3.18 中情形 (i) 的饱和正交表可以构造许多均匀设计.

3.6.4 进一步结果

在上面的小节中, 主要讨论了离散偏差 (2.52) 下均匀设计的构造, 且生成的设计大部分达到定理 2.22 给出的下界. 基于定理 2.23, 将这些设计删掉任意一列或添加任一平衡列, 仍然得到一个均匀设计, 因为这样得到的设计的任两行之间的相遇数最多差 1. 由 (2.52), 可以看出这种偏差将不同水平数的因子视为同等重要的. 而离散偏差 (2.52) 和投影偏差 (6.25) 给不同水平数的因子赋予了不同权重, 接下来, 将通过取 (6.25) 中的 $j = 2$ 且 $\gamma = 1$, 给出该投影偏差下的二维投影均匀设计的一些结果.

注意到, 上面讨论的所有均匀设计 $U_n(q^s)$ 都是基于 U-型设计而构造的, 因而要求试验次数 n 是因子水平数 q 的倍数. 在本小节中, 也会给出 n 不能被 q 整除情形下近似 U-型设计和均匀设计的某些结果.

A. 二维投影均匀设计

对一个有 n 次试验, s 个因子以及权重 $w_k>0$ 的因子设计 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$, 令

$$\delta_{ij}^{\mathrm{w}}(\mathcal{P})=\sum_{k=1}^{s}w_k\delta_{x_{ik}x_{jk}}. \tag{3.18}$$

显然, $\delta_{ij}^{\mathrm{w}}(\mathcal{P})$ 是 (2.48) 的一个推广并称之为行 $\boldsymbol{x}_i$ 与 $\boldsymbol{x}_j$**加权相遇数** (weighted coincidence number). 令

$$\delta=\sum_{k=1}^{s}w_k(n/q_k-1)/(n-1),$$

$$\Delta=\left\{\sum_{k=1}^{s}w_k\delta^{(k)}:\delta^{(k)}=0,1,\ k=1,\cdots,m\right\}.$$

记 δ_L 及 δ_U 为 Δ 中满足 $\delta_L\leqslant\delta<\delta_U$ 且离 δ 最近的两个数. 那么, 基于优势理论 (Marshall and Olkin, 1979) 以及基于 (6.25) 的投影均匀性与其他准则 (即广义最小低阶混杂、最小低阶矩混杂及最小 χ^2 准则) 的关系, Liu 等 (2006) 得到了下面的二维投影偏差的下界.

定理 3.20 假定权重 $w_k=\lambda q_k$, $1\leqslant k\leqslant m$, 则对任一 $U(n;q_1\cdots q_s)$ 设计 $\boldsymbol{U}$, 当 (6.25) 式取 $\gamma=1$ 时, $[D_{(2)}(\boldsymbol{U};\mathcal{K})]^2$ 满足

$$[D_{(2)}(\boldsymbol{U};\mathcal{K})]^2\geqslant\frac{n-1}{2n\lambda^2}\Big\{(\delta_U+\delta_L)\delta-\delta_U\delta_L\Big\}$$
$$-\frac{1}{2}\left\{s(s-1)+\sum_{k=1}^{s}q_k-\frac{1}{n}\left(\sum_{k=1}^{s}q_k\right)^2\right\}.$$

等式成立当且仅当对任一 i, 在 $(n-1)$ 个值 $\delta_{1i}^{\mathrm{w}}(\boldsymbol{U}),\cdots,\ \delta_{(i-1)i}^{\mathrm{w}}(\boldsymbol{U}),\delta_{i(i+1)}^{\mathrm{w}}(\boldsymbol{U}),\cdots,$ $\delta_{in}^{\mathrm{w}}(\boldsymbol{U})$ 中, 有 $(n-1)\dfrac{\delta_U-\delta}{\delta_U-\delta_L}$ 个取 δ_L, 且 $(n-1)\dfrac{\delta-\delta_L}{\delta_U-\delta_L}$ 个取 δ_U.

值得注意的是, 达到上述下界的大部分设计都是超饱和设计, 且 3.6.2 节得到的所有任意两行的相遇数为常数的对称设计在 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下也是最优的.

定理 3.20 给出了下界能够达到的一个条件. 对于 $(n,s,q_1,\cdots,q_s)$ 的某些取值, 相应的下界可以达到. 比如, 当 $\boldsymbol{U}$ 是一个饱和 $\mathrm{OA}(n;q_1,\cdots,q_s;2)$ 时, $\boldsymbol{\delta}(\boldsymbol{U})=$

$(\lambda(s-1),\cdots,\lambda(s-1))$(Mukerjee and Wu, 1995), 则下界可达到. 基于这一事实, Li 等 (2004) 将算法 3.8 的 FSOA 方法进行了推广, 来构造 $\chi^2(\boldsymbol{U})$ 最优的非对称超饱和设计并研究了所得设计的性质. 基于 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 与 $\chi^2(\boldsymbol{U})$ 的联系 (见 (Liu et al., 2006)) 由他们的方法构造的设计在 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下也是均匀设计.

Fang 等 (2004a) 是考虑非对称超饱和设计构造的另一篇论文, 在 3.6.2 节刚介绍过. 由 (Fang et al., 2004a) 的总结可知, 他们得到的设计的任意两行相遇数恰为 1. 同样可以看出他们得到的设计中大部分都有形式 $U(n;p^1q^{s-1})$. 对于 $U(n;p^1q^{s-1})$ 设计, Liu 等 (2006) 给出了如下结果.

定理 3.21 令 $\boldsymbol{U}$ 为一个 $U(n;p^1q^{s-1})$ 设计, 其中 $p\leqslant q$ 且 $n/p+(s-1)n/q-s=n-1$. 如果 $\boldsymbol{U}$ 的任意两行相遇数恰为 *1*, 那么 $\boldsymbol{U}$ 在 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下是最优的.

故由该定理可以得出结论: 3.6.2 小节得到的均匀设计 $U_n(p^1q^{s-1})$(其中 $p\leqslant q$) 在 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下也是均匀的.

列并置 (column juxtaposition) 方法也可以用来构造非对称均匀设计. 由 (Li et al., 2004) 及 (Liu et al., 2006), 有如下推论.

推论 3.22 令 $\boldsymbol{U}_k$ $(1\leqslant k\leqslant l)$ 是一组试验数相同的平衡设计. 如果取自然权重 $w_k=\lambda q_k$, $k=1,\cdots,m$, 且每个设计 $\boldsymbol{U}_k$ 的行之间的加权相遇数是常数, 则 $\boldsymbol{U}=(\boldsymbol{U}_1,\cdots,\boldsymbol{U}_l)$ 在 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下是最优的.

基于该推论, 不仅可以由强度为 2 的饱和正交表构造 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下的均匀设计, 也可以由符合推论中给定条件的设计来构造 $D_{(2)}(\boldsymbol{U};\mathcal{K})$ 下的均匀设计, 比如, 由 (Liu and Zhang, 2000; Fang et al., 2000a) 中的设计, 以及 3.6.2 小节和 3.6.3 小节中介绍的设计.

B. 近似 U-型设计和均匀设计

当 n 不能被 q 整除时, 设 $n=qt+r(0<r<q)$. 在这种情形, 安排设计 $\boldsymbol{U}$, 使得每列中有 $q-r$ 个水平各出现 t 次, 其余 r 个水平各出现 $t+1$ 次. 这样可以保证 $\boldsymbol{U}$ 的每列中不同水平出现的次数尽可能相同. 这种设计被称为近似 U-型设计 (nearly U-type design), 并记为 $\mathrm{NU}(n;q^s)$(Fang et al., 2004d). 为完整起见, 加入 $r=0$ 的情况. $n=qt$ 时的 $\mathrm{NU}(n;q^s)$ 就是一个 U-型设计.

对离散偏差 (2.53), Fang 等 (2004d) 得到了如下下界, 它推广了定理 2.22 中给出的下界.

定理 3.23 令 n, s 和 q 为正整数且$n=qt+r$, $0\leqslant r\leqslant q-1$. 记 $\boldsymbol{U}$ 是一个有 n 次试验, s 个 q 水平因子的部分因子设计. 设 $u=\dfrac{st(n-q+r)}{n(n-1)}$ 且 $\lambda=\lfloor u\rfloor$. 那么对于离散偏差 (2.53), 有

$$[\mathrm{DD}(\boldsymbol{U})]^2 \geqslant -\left(\frac{a+(q-1)b}{q}\right)^s + \frac{a^s}{n} + \frac{b^s(n-1)}{n}\left[(\lambda+1-\mu)\left(\frac{a}{b}\right)^{\lambda} + (\mu-\lambda)\left(\frac{a}{b}\right)^{\lambda+1}\right], \tag{3.19}$$

并且 (3.19) 式右边 $\mathrm{DD}(\boldsymbol{U})$ 的下界能够达到当且仅当对任意 $i \neq j$, (2.48) 式中定义的 $\delta_{ij}(\boldsymbol{U})$ 取同一值 λ 或只取两个值 λ 和 $\lambda+1$.

现在称一个 $\mathrm{NU}(n;q^s)$ 是在 $\mathrm{DD}(\boldsymbol{U})$ 下的均匀设计, 也记为 $U_n(q^s)$, 如果它的偏差 $\mathrm{DD}(\boldsymbol{U})$ 达到了所有这样 $\mathrm{NU}(n;q^s)$ 中的最小值. 显然, $[\mathrm{DD}(\boldsymbol{U})]^2$ 值等于 (3.19) 式所给下界的一个设计 $\boldsymbol{U}$ 是一个均匀设计.

Fang 等 (2004d) 进一步探索了均匀设计与组合设计理论中的可分解填充设计及覆盖设计之间的关系. 利用平行类不同的可分解填充设计及覆盖设计, 可生成许多新的均匀设计的无穷类. 更多的细节可参见 (Fang et al., 2004d).

习　　题

3.1　对于参数 $n=8, s=2$, 给出设计空间 $\mathcal{D}(8;C^2), \mathcal{D}(8;8^2), \mathcal{U}(8;8^2)$ 和 $\mathcal{U}(8;2\times 4)$.

3.2　把例 3.2 的设计 $\boldsymbol{U}$ 分别变化到 C^4 中的 $\boldsymbol{X}_{\mathrm{lft}}, \boldsymbol{X}_{\mathrm{ctr}}, \boldsymbol{X}_{\mathrm{ext}}$ 和 $\boldsymbol{X}_{\mathrm{mis}}$. 分别计算 C^4 中的这四个设计的可卷偏差、中心化偏差和混合偏差. 判断这些设计是否等价.

3.3　给出能用好格子点法构造均匀设计表 $U_n(n^s)$ 的必要条件.

3.4　用 3.3.1 小节中介绍的 Fibonacci 序列的方法构造 $U(n;n^2)$, $n=5,8,13$, 并比较其与表 3.1 和表 3.2 相应设计的中心化偏差.

3.5　设某试验有三个因子，包括温度、压力和反应时间，其取值分别为 [50, 100]°C, [3, 5]atm 和 [10, 25]min. 则试验区域为 $\mathcal{X}=[50,80]\times[3,6]\times[10,25]$. 设每个因子有 4 个水平. 两个可能的设计 $\boldsymbol{Z}_1$ 和 $\boldsymbol{Z}_2$, 以及映射到 $[0,1]^2$ 中的矩阵分别如下所示.

$$\boldsymbol{Z}_1 = \begin{bmatrix} 60 & 6 & 25 \\ 70 & 6 & 15 \\ 60 & 3 & 15 \\ 80 & 5 & 20 \\ 50 & 5 & 10 \\ 70 & 3 & 25 \\ 80 & 4 & 10 \\ 50 & 4 & 20 \end{bmatrix} \Rightarrow \boldsymbol{X}_1 = \begin{bmatrix} 0.3333 & 1 & 1 \\ 0.6667 & 1 & 0.3333 \\ 0.3333 & 0 & 0.3333 \\ 1 & 0.6667 & 0.6667 \\ 0 & 0.6667 & 0 \\ 0.6667 & 0 & 1 \\ 1 & 0.3333 & 0 \\ 0 & 0.3333 & 0.6667 \end{bmatrix},$$

$$
\boldsymbol{Z}_2 = \begin{bmatrix} 61.25 & 5.625 & 23.125 \\ 68.75 & 5.625 & 15.625 \\ 61.25 & 3.375 & 15.625 \\ 76.25 & 4.875 & 19.375 \\ 53.75 & 4.875 & 11.875 \\ 68.75 & 3.375 & 23.125 \\ 76.25 & 4.125 & 11.875 \\ 53.75 & 4.125 & 19.375 \end{bmatrix} \Rightarrow \boldsymbol{X}_2 = \begin{bmatrix} 0.375 & 0.875 & 0.875 \\ 0.625 & 0.875 & 0.375 \\ 0.375 & 0.125 & 0.375 \\ 0.875 & 0.625 & 0.625 \\ 0.125 & 0.625 & 0.125 \\ 0.625 & 0.125 & 0.875 \\ 0.875 & 0.375 & 0.125 \\ 0.125 & 0.375 & 0.625 \end{bmatrix}.
$$

回答如下问题:

(1) 比较设计 $\boldsymbol{Z}_1$ 和 $\boldsymbol{Z}_2$，并给出你的建议.

(2) 设计 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$，一个有端点 0 和 1，另一个没有. 给出相应的映射.

3.6 计算欧拉函数 $\phi(n)$, $n = 8, 9, 10, 14, 15$，并给出相应的生成向量候选集 $\mathcal{H}_n$.

3.7 给出 3 个 $\phi(n) > \phi(n+1)$ 的情形，并比较删行好格子点法和好格子点法的性能.

3.8 给出候选集 $\mathcal{G}_{8,2}$. 在混合偏差下，把 $\mathcal{G}_{8,2}$ 中的设计分类，使得类内的设计是等价的，而不同类的设计有不同的混合偏差.

3.9 给定参数 (n, s)，寻找候选集 $\mathcal{A}_{n,s}$, $n = 11, s = 2, 3, 4, 5$, 且 $n = 30$, $s = 2, 3, 4, 5$, 其中

$$
\mathcal{A}_{n,s} = \{a : a < n,\ \gcd(a^j, n) = 1,\ j = 1, \cdots, s-1; 1, a, a^2, \cdots, a^{s-1} (\widetilde{\text{mod}}\ n) \text{互不相同}\}.
$$

3.10 从均匀设计 $U_{47}(47^3)$ 出发，用切割法构造在混合偏差意义下的近似均匀设计 $U_8(8^3)$.

3.11 设 $n = 47$ 和 $s = 46$. 计算相应好格子点 U 的 L_1-距离和混合偏差. 考虑简单线性水平置换 $U + i\ (\widetilde{\text{mod}}\ n)$ 并给出你的结论.

3.12 一个 n 阶的**拉丁方设计**是一个 $n \times n$ 的矩阵，其元素来自 n 个不同的元素，每个元素在每行和每列都只出现一次.

(1) 分别给出一个 3 阶和一个 4 阶拉丁方设计.

(2) 搜索文献中关于正交拉丁方设计的定义.

(3) 给出一种由正交拉丁方设计构造正交设计 $L_{n^2}(n^{n-1})$ 的方法.

(4) 数独是一种特殊的拉丁方设计. 考虑下面的 9×9 的数独并填充其中空格处数字:

9		4	2			8	3	
	5	7	4		9			
6	2			7	1		9	5
8	9	2	7	3			1	4
1			6		4	9		
4	7				8		5	
5				4		3	8	2
	8		5					9
			8	9	2		6	1

第 4 章　构造均匀设计的优化算法

第 3 章介绍了构造均匀设计的确定性方法，包括好格子点法及其推广方法等. 一般地, 均匀设计网站中大部分的均匀设计都是由数值搜索算法得到的, 其搜索空间往往为 $\mathcal{U}(n;n^s)$ 或 $\mathcal{U}(n;q_1\times\cdots\times q_s)$. 本章将介绍搜索均匀设计的一些高效优化方法. 4.2 节和 4.3 节分别介绍门限接受法和整数规划法这两个算法.

4.1　均匀设计的数值搜索法

由均匀设计的定义 3.2 可知, 给定均匀性测度 D, 构造一个 n 个试验点的均匀设计, 即为寻找一个设计 $\mathcal{P}^*\in\mathcal{D}(n;C^s)$ 使得

$$D(\mathcal{P}^*)=\min_{\mathcal{P}\in\mathcal{D}(n;C^s)}D(\mathcal{P}), \tag{4.1}$$

其中 $\mathcal{D}(n;C^s)=\{\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}:\boldsymbol{x}_i\in C^s, i=1,\cdots,n\}$. 易知, 牛顿法等经典的优化算法并不适用于这类问题的求解.

首先给出关于数学优化问题的一些概念. 设 $f(\boldsymbol{x})=f(x_1,\cdots,x_m)$ 为集合 $\mathcal{X}$ 上的连续函数. 我们希望找到一个 $\boldsymbol{x}^*\in\mathcal{X}$, 使得

$$f(\boldsymbol{x}^*)=\min_{\boldsymbol{x}\in\mathcal{X}}f(\boldsymbol{x}), \tag{4.2}$$

其中 $\boldsymbol{x}^*$ 被称为*全局最小点*且 $f(\boldsymbol{x}^*)$ 为 $\mathcal{X}$ 中具有最小 f 值. 称函数 f 为*目标函数*或*损失函数*. 一般地, $\mathcal{X}$ 是欧氏空间 R^m 的子集, 称其为*搜索域*或*搜索空间*; 称 $\mathcal{X}$ 中的点为*候选解*或*可行解*. 通常 $\mathcal{X}$ 中的解需要满足一些约束条件. 一个*局部最小解* $\boldsymbol{x}_{\rm local}\in\mathcal{X}$ 满足

$$f(\boldsymbol{x}_{\rm local})\leqslant f(\boldsymbol{x}), \text{任意 } \boldsymbol{x} \text{ 满足 } ||\boldsymbol{x}-\boldsymbol{x}_{\rm local}||\leqslant\delta \text{且} \delta>0.$$

当搜索域 $\mathcal{X}$ 中仅有一个局部最小解时, 则其也是全局最小解, 此时有很多高效算法可以找到该最优解, 其中*牛顿法*是最常用的方法. 该方法从某个初始解 $\boldsymbol{x}_0\in\mathcal{X}$ 开始, 然后迭代地变化变量取值以最小化目标函数. 记 $\boldsymbol{x}_c=\boldsymbol{x}_0$. 每步迭代都需要计算目标函数 f 在当前变量值下的导数值, 然后确定一个更新点 $\boldsymbol{x}_{\rm new}\in\mathcal{X}$. 若 $f(\boldsymbol{x}_{\rm new})<f(\boldsymbol{x}_c)$, 用 $\boldsymbol{x}_{\rm new}$ 代替 $\boldsymbol{x}_c$ 并进行下一步迭代. 当 $\boldsymbol{x}_{\rm new}=\boldsymbol{x}_c$ 时, 算法中止. 然而, 当目标函数存在多个局部最小值时, 牛顿法往往只能收敛到局部最优解. 搜

索均匀设计的优化问题 (3.6) 通常存在很多局部最小解. 对于任一优化问题, 需要估计该问题的计算复杂度.

为了得到一个满足 (4.1) 式的均匀设计, 需要考虑下列的困难:

(1) *高维问题*. 这本质上是一个 ns 维的优化问题 (见 3.2.1 小节). 因此, 即使 n 和 s 不太大时, 我们也面临一个高维数的最小化问题.

(2) *候选集*. 试验域上存在无穷多个候选集.

(3) *多个局部最小解*. 目标函数存在多个局部最小解.

由于存在这些困难, 搜索一个均匀设计 $U_n(n^s)$ 是一个 NP-*难问题*, 其中 NP 表示 "非确定多项式时间". 我们知道, 不存在一个多项式时间的算法可以解决 NP-难问题. 因此, 我们可以减小设计空间以降低复杂度, 例如, U-型设计集 $\mathcal{U}(n;n^s)$, 或更一般的 $\mathcal{U}(n;q_1\times\cdots\times q_s)$, 其中的每个设计都有 n 行、s 列, 其水平分别为 $q_1,\cdots,q_s$, 且每列的各水平出现次数一样多. 为了简单起见, 把 U-型设计集 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 记为 $\mathcal{U}$. 则 (4.1) 式的构造均匀设计的优化问题变为寻找一个设计 $\mathcal{P}^*\in\mathcal{U}$ 使得

$$D(\mathcal{P}^*)=\min_{\mathcal{P}\in\mathcal{U}}D(\mathcal{P}), \tag{4.3}$$

这里, 偏差可以选取为中心化偏差、可卷偏差、混合偏差或其他类型偏差. 然而, 求解最小化问题 (4.3) 仍是一个有挑战性的问题, 因为其还存在下面的困难:

(1) 设计空间不是一个紧区域. 此时, 经典的优化算法都失效了, 因为不存在连续性和导数的概念. 因此, 我们需要现代优化技术, 如随机优化算法等.

(2) 随着 n 和 s 的增大, 设计空间中的候选集个数呈指数增长. 在 $\mathcal{U}$ 中搜索一个均匀设计是一个 NP-难问题.

(3) 任意置换行次序和列次序, 并不改变偏差值. 因此均匀设计不唯一，存在多个局部最小解，其导致难以确定算法的终止准则.

(4) 当一个设计 $\boldsymbol{x}_{\text{new}}$ 达到 2.6 节中的目标函数的下界时, 算法终止. 然而, 在很多情形下, 确定偏差的紧下界仍是一个公开问题. 一个紧下界是指存在某设计达到该下界.

一个好的搜索均匀设计的算法需面临上述挑战. 为此, 文献中提出了许多随机优化算法来构造均匀设计, 如*模拟退火算法*(Kirkpatrick et al., 1983; Morris and Mitchell, 1995)、随机进化算法 (Jin et al., 2005)、门限接受法 (Dueck and Scheuer, 1990; Winker and Fang, 1998; Fang et al., 2002d, 2003b, 2005, 2006b) 以及整数规划法 (Fang and Ma, 2001b; Zhou et al., 2013) 等.

构造均匀设计的随机优化算法的一般过程为: 选择一个初始设计 $\boldsymbol{U}_0$ 作为当前设计 $\boldsymbol{U}_c$; 在 $\boldsymbol{U}_c$ 的邻域中选择另一个设计 $\boldsymbol{U}_{\text{new}}$, 若满足接受准则, 则把当前设计 $\boldsymbol{U}_c$ 替换为 $\boldsymbol{U}_{\text{new}}$; 持续该过程直到满足终止准则. 因此, 一个随机优化算法有下面的一般框架.

算法 4.1 (构造均匀设计的框架)

步骤 0 *输入*. 给定试验次数 n、因子个数 s、设计空间 $\mathcal{U}$ 和作为优化目标函数的偏差准则 $D(\cdot)$.

步骤 1 *初始解*. 选择一个初始解 $\boldsymbol{U}_0 \in \mathcal{U}$. 设当前设计 $\boldsymbol{U}_c = \boldsymbol{U}_0$.

步骤 2 *评价和迭代*. 在 $\boldsymbol{U}_c$ 的邻域中选择一个新设计, 记为 $\boldsymbol{U}_{\rm new}$. 比较 $\boldsymbol{U}_{\rm new}$ 和 $\boldsymbol{U}_c$ 的偏差值并决定是否把当前设计 $\boldsymbol{U}_c$ 替换为 $\boldsymbol{U}_{\rm new}$. 若需要替换, 则重复该步骤, 否则根据给定的跳跃准则把当前设计变到某新设计.

步骤 3 *终止准则*. 若跳跃准则已结束, 则当前设计 $\boldsymbol{U}_c \in \mathcal{U}$ 已有足够好的均匀性, 则输出该设计并结束算法.

该框架中的每一步骤都可以有不同的实现方式, 其取决于设计空间的复杂度, 选定偏差的结构, 设计的质量和计算时间. 例如, 设计空间可取为 n 行、s 列的全体设计集或其子集, 偏差可以选取为可卷偏差、中心化偏差、混合偏差或其他偏差. 初始设计可从设计空间中随机选取一个设计或选择一个具有较好均匀性的设计. 我们也有多种方法定义某设计在 $\mathcal{U}$ 中的邻域, 另外也有从局部最小解中跳跃出来的不同跳跃准则. 后面将讨论每一步骤的不同实现方式. 为了方便起见, 我们可能在后续算法中重新调整算法 4.1 中各步骤的顺序.

此外, 算法 4.1 中, 选定偏差的下界将很有用, 因为当迭代过程中某设计达到下界时, 则直接终止迭代过程. 不同偏差的下界参见 2.6 节.

4.2 门限接受法

门限接受法 (threshold accepting, TA) 是一种构造均匀设计的有效优化算法. 该方法最早由 Dueck 和 Scheuer (1990) 提出. Winker 和 Fang (1997) 首先把 TA 算法用于伪蒙特卡罗法, 并在星偏差意义下构造均匀设计. 后面他们获得具有更低星偏差的 U-型设计 (Winker and Fang, 1998). Fang 等 (2000b) 在不同偏差下应用 TA 算法寻优, 并指出很多已有的正交设计可以在中心化偏差或可卷偏差意义下应用 TA 算法而得到. Fang 等 (2002d) 研究拉丁超立方体抽样并给出搜索均匀设计的数值方法的一些建议, 基于这些信息在中心化偏差意义下应用 TA 算法构造均匀设计. Fang 等 (2003b) 进一步应用 TA 算法构造在中心化偏差和可卷偏差下的均匀设计. 还有很多其他作者用 TA 算法构造均匀设计, 例如, Fang 等 (2005), Fang 等 (2006b) 与 Zhou 和 Fang (2013), 其中 Zhou 和 Fang (2013) 考虑用 TA 算法构造大均匀设计. TA 算法可以应用于离散优化问题. 本节将介绍 TA 算法的详细过程.

在构造均匀设计的 TA 算法中, 需要考虑下面的各方面:

目标函数 TA 算法考虑下面的优化问题

$$\min_{\boldsymbol{U}\in\mathcal{U}} D(\boldsymbol{U}), \tag{4.4}$$

其中 $\mathcal{U}$ 为所有 U-型设计 $\mathcal{U}(n;q_1\times\cdots\times q_s)$ 的子集, D 为某种均匀性测度, 例如, 中心化偏差或混合偏差等.

图 4.1 给出在设计空间 $\mathcal{U}$ 上构造均匀设计的门限接受法的示意图. 前面提到的文献中也提出了一些改进方法. 一个设计 $\mathcal{P}\in\mathcal{U}$ 可以表示为 $n\times s$ 的矩阵 $\boldsymbol{U}=(u_{ij}), u_{ij}\in[0,1]$. 为了方便起见, 本书中 $\boldsymbol{U}$ 和 $\mathcal{P}$ 都用于表示某个设计. TA 算法中的一些具体的步骤解释如下.

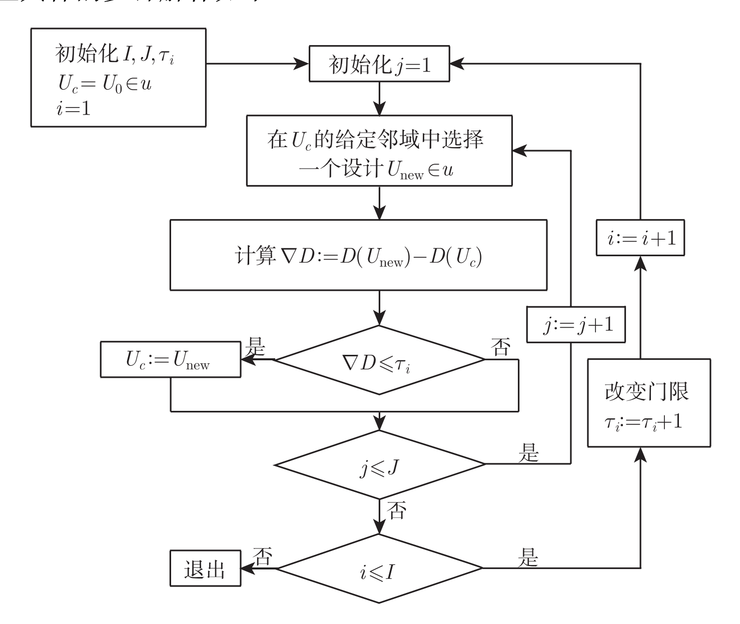


图 4.1 搜索均匀设计的门限接受法

初始设计 初始设计 $\boldsymbol{U}_0$ 可以从设计空间 $\mathcal{U}$ 中随机选一个设计而得到, 或者用好格子点法得到, 或其他已知的低偏差设计.

给定设计空间 $\mathcal{U}(n;q^s)$ 或 $\mathcal{U}(n;q_1\times\cdots\times q_s)$, $s<n$, 为了得到一个低偏差的初始设计, 可以先选择一个低偏差设计 $U(n;n^s)$, 然后用下面的伪水平变换方法把它变化到对称设计 $U(n;q^s)$ 或非对称设计 $U(n;q_1\times\cdots\times q_s)$(Fang et al., 2006):

$$\{(i-1)n/q+1,\cdots,in/q\}\to i,\quad i=1,\cdots,q, \tag{4.5a}$$

或

$$\{(i-1)r_j+1,\cdots,ir_j\}\to i,\quad i=1,\cdots,q_j, \tag{4.5b}$$

其中 q_j 为第 j 个因子的水平数, $r_j = n/q_j$.

门限和置换规则　门限值 τ 是一个在迭代中逐渐变小的非负数. 经典的局部搜索算法使得目标函数越来越小, 且当 $\nabla D = D(\boldsymbol{U}_{\rm new}) - D(\boldsymbol{U}_c) \leqslant 0$ 时, 把当前设计 $\boldsymbol{U}_c$ 替换为 $\boldsymbol{U}_{\rm new}$. 局部搜索算法容易陷入局部最小偏差设计, 其可能离全局最小偏差设计有较远的距离. TA 算法用另一个规则: 当 $\nabla D = D(\boldsymbol{U}_{\rm new}) - D(\boldsymbol{U}_c) \leqslant \tau$ 时, 当前设计 $\boldsymbol{U}_c$ 被替换为 $\boldsymbol{U}_{\rm new}$, 其中 $\tau \geqslant 0$. 当 $\tau > 0$ 时, 迭代算法存在从局部最小解跳出来的机会. 当 τ 不太小时, 迭代过程像随机游走; 当 τ 变小时, 迭代过程可以在 $\boldsymbol{U}_c$ 的小邻域中移动; 当 $\tau = 0$, 迭代过程终止于 $\boldsymbol{U}_c$ 的邻域中的局部最小解. $\boldsymbol{U} \in \mathcal{U}$ 的邻域需要事先定义, 我们将在后面讨论.

门限值将严重影响迭代结果. 文献中, 可以选择一个合适的正整数 I, 并确定一系列门限 $\tau_1 > \tau_2 > \cdots > \tau_{I-1} > \tau_I = 0$. 在第 i 次外循环中, 当 $\nabla D > \tau_i$ 和 $j < J$ 时, 在 $\boldsymbol{U}_c$ 的邻域中随机选择一个新设计, 并令 $j := j + 1$. 一般地, 外循环的次数 I 需远小于内循环的次数 J. 图 4.1 给出 I 和 J 的作用. I 和 J 的选择依赖于设计的参数. 当 $n \leqslant 1000$ 时, 文献中建议 $I \in [10, 100]$, $J \in [10^4, 10^5]$, 且随着 n 的增大, 需要增加 I 和 J 的取值 (Winker and Fang, 1997; Zhou and Fang, 2013). 例如, 当 $n \leqslant 100$ 时, 可取 $I = 10, J = 10^5$; 当 $n \geqslant 300$ 时, $I = 100,\ J = 10^5$. 为了获得更合适的 I 和 J, 可以事先做一些经验的比较.

为了确定门限值 $[\tau_1, \cdots, \tau_I]$, 可以随机地从 $\mathcal{U}$ 中选择 M 个设计, 分别计算其目标函数值, 从而得到这 M 个取值的顺序统计量. 记目标函数的最大值和最小值分别为 $D_{\max}$ 和 $D_{\min}$, 则取值范围 $R = D_{\max} - D_{\min}$. 可以选取第一个门限 $\tau_1 = D_{\min} + \alpha R\ (0 < \alpha < 1)$. 剩余的 $I - 1$ 的门限可以迭代地确定 $\tau_i = f(\tau_{i-1})$, 其中 $f(\cdot)$ 可以取为线性函数或其他函数, 例如

$$\tau_i = \frac{I - i}{I}\tau_{i-1}, \quad i = 2, \cdots, I. \tag{4.6}$$

我们可以事先选择不同的 α 值, 并比较不同的性能, 最终选择一个最佳的 α.

另一种确定 τ_1 的方法如下. 首先随机选择一个设计 $\boldsymbol{U} \in \mathcal{U}$, 然后在 $\boldsymbol{U}$ 的邻域中随机产生 N 个设计 $\boldsymbol{U}_j$. 计算这 N 个差异 $\nabla D_j = D(\boldsymbol{U}_j) - D(\boldsymbol{U})$, 并得到 ∇D_j 中所有取值非负的经验分布, 记为 F_1. τ_1 可以取为 F_1 的 t 分位数, 例如, $t = 5$.

邻域　由于设计空间为有限集, 目标函数并不存在连续性, 因此这里邻域的定义不能用连续的概念. 若两个设计在同一邻域中, 我们希望它们的设计结构相似. 设 $\boldsymbol{U}_c$ 为 $\mathcal{U}$ 中的某设计矩阵, 定义其邻域 $\mathcal{N}(\boldsymbol{U}_c)$ 时需要考虑几个方面. 首先, 正如之前已指出的, $\mathcal{N}(\boldsymbol{U}_c) \subset \mathcal{U}$ 且每个 $\boldsymbol{U}_{\rm new} \in \mathcal{N}(\boldsymbol{U}_c)$ 应为一个 U-型设计. 其次, 从邻域的概念考虑, 设计 $\boldsymbol{U}_c$ 和其邻域 $\mathcal{N}(\boldsymbol{U}_c)$ 中的设计应该相差不大. 再次, 算法的计算复杂度在很大程度上依赖于计算 ∇D, 即从 $\boldsymbol{U}_c$ 变到 $\boldsymbol{U}_{\rm new}$ 时目标函数的变化情况. 因此, 若计算 ∇D 时不需要直接计算 $\boldsymbol{U}_{\rm new}$ 的目标函数值, 而通过某些快速

方法得到其取值, 则可以大大提高算法的速度.

为了满足这三个要求, 我们可以随机选取 $\boldsymbol{U}_c$ 的一列或几列, 并在所选列中各随机选择两个元素并相互交换, 从而得到一个新设计, 把全部的这些设计定义为 $\boldsymbol{U}_c$ 的邻域. 例如

$$\mathcal{N}_1(\boldsymbol{U}_c) = \{\boldsymbol{U} : \text{置换}\boldsymbol{U}_c\text{的某一列中的两个元素}\},$$

或定义一个更大的邻域如下:

$$\mathcal{N}_2(\boldsymbol{U}_c) = \{\boldsymbol{U} : \text{置换}\boldsymbol{U}_c\text{的某两列中各两个元素}\}.$$

Winker 和 Fang (1997) 给出选择邻域大小的一些讨论. 定义的邻域越大, 我们可以越自由地选择 $\boldsymbol{U}_{\text{new}}$.

实际上, 在大多数 TA 算法中, $\boldsymbol{U}_c$ 的邻域是其一个小的扰动. 前面定义的 $\mathcal{N}_1(\boldsymbol{U}_c)$ 和 $\mathcal{N}_2(\boldsymbol{U}_c)$ 中某一列或两列是随机选择的. Fang 等 (2005) 提出另外两种方法来选择一个新设计.

(1) *“极大和极小距离对” 法*. 设 $(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2})$ 和 $(\boldsymbol{x}_{j1}, \boldsymbol{x}_{j2})$ 分别为当前设计 $\boldsymbol{U}_c$ 中具有最大距离的行对和具有最小距离的行对. 这里的距离可定义为 L_p-距离或其他距离. 从 $\boldsymbol{x}_{i_1}$ 和 $\boldsymbol{x}_{i_2}$ 这两行中随机选择一行, 记为 $\boldsymbol{x}_i$; 并从 $\boldsymbol{x}_{j_1}$ 和 $\boldsymbol{x}_{j_2}$ 这两行中随机选择一行, 记为 $\boldsymbol{x}_j$. 然后随机选择第 k 列, 若行向量 $\boldsymbol{x}_i$ 中 x_{ik} 不等于行向量 $\boldsymbol{x}_j$ 中的元素 x_{jk}, 则置换 x_{ik} 和 x_{jk} 以产生新设计 $\boldsymbol{U}_{\text{new}}$.

(2) *“单行最大总距离和最小总距离” 法*. 记 $d_{ij} = d(\boldsymbol{x}_i, \boldsymbol{x}_j)$, 其中 $d(\cdot,\cdot)$ 为事先确定的距离准则. 记 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_j$ 分别是当前设计 $\boldsymbol{U}_c$ 中具有最大总距离和最小总距离的行向量, 即对所有的 k 计算 $\sum_{t\neq k} d_{tk}, k = 1, \cdots, s$, 其中 $\sum_{t\neq i} d_{ti}$ 取值最大, 而 $\sum_{t\neq j} d_{tj}$ 最小. 随机选择第 k 列, 若行向量 $\boldsymbol{x}_i$ 中 x_{ik} 不等于行向量 $\boldsymbol{x}_j$ 中的元素 x_{jk}, 则置换 x_{ik} 和 x_{jk} 以产生新设计 $\boldsymbol{U}_{\text{new}}$.

每一个方法都有自己的优点. 与 “极大和极小距离对” 法相比, “单行最大总距离和最小总距离” 法可能收敛速度更快, 不过前者有更多的机会跳出局部最小解. 这两种方法中采用随机选取方法的主要目的都是加快收敛速度且能有机会跳出局部最优解.

迭代公式 在 TA 算法中用偏差来区分不同设计的均匀性. 在每步迭代时需要计算 $D(\boldsymbol{U}_{\text{new}})$. 因为 $D(\boldsymbol{U}_{\text{new}})$ 在邻域 $\mathcal{N}(\boldsymbol{U}_c)$ 中, $D(\boldsymbol{U}_{\text{new}})$ 和 $D(\boldsymbol{U}_c)$ 的计算式之间应当存在紧密联系. 若能得到 $D(\boldsymbol{U}_{\text{new}})$ 和 $D(\boldsymbol{U}_c)$ 之间的迭代公式, 则可以大大减少计算 $D(\boldsymbol{U}_{\text{new}})$ 的时间. 一般地, 对于不同偏差存在不同的迭代公式.

A. *中心化偏差的迭代公式*

Jin 等 (2005) 给出一个计算中心化偏差的简单方法. 设 $\boldsymbol{U}_c = (u_{ik})$ 为当前设计, 其导出设计矩阵为 $\boldsymbol{X} = (x_{ik})$, $\boldsymbol{Z} = (z_{ik})$ 为中心化的设计矩阵, 其中 $z_{ik} = x_{ik} - 0.5$.

设 $\boldsymbol{C}=(c_{ij})$ 为对称矩阵, 其元素为

$$c_{ij}=\begin{cases}\dfrac{1}{n^2}\prod_{k=1}^{s}\dfrac{1}{2}(2+|z_{ik}|+|z_{jk}|-|z_{ik}-z_{jk}|), & i\neq j,\\[2ex] \dfrac{1}{n^2}\prod_{k=1}^{s}(1+|z_{ik}|)-\dfrac{2}{n}\prod_{k=1}^{s}\left(1+\dfrac{1}{2}|z_{ik}|-\dfrac{1}{2}z_{ik}^2\right), & \text{否则}.\end{cases}$$

可以证明

$$[\mathrm{CD}(\boldsymbol{U}_c)]^2=\left(\frac{13}{12}\right)^2+\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}.$$

对任意 $1\leqslant t\leqslant n$ 和 $t\neq i,j$, 设

$$\gamma(i,j,k,t)=\frac{2+|z_{jk}|+|z_{tk}|-|z_{jk}-z_{tk}|}{2+|z_{ik}|+|z_{tk}|-|z_{ik}-z_{tk}|}.$$

交换 x_{ik} 和 x_{jk} 后, 新设计 $\boldsymbol{U}_{\text{new}}$ 的平方中心化偏差为

$$[\mathrm{CD}(\boldsymbol{U}_{\text{new}})]^2=[\mathrm{CD}(\boldsymbol{U}_c)]^2+\nabla(\mathrm{CD}),$$

其中

$$\nabla(\mathrm{CD})=c_{ii}^*-c_{ii}+c_{jj}^*-c_{jj}+2\sum_{t=1,t\neq i,j}^{n}(c_{it}^*-c_{it}+c_{jt}^*-c_{jt}),$$

且 $c_{it}^*=\gamma(i,j,k,t)c_{it}, c_{jt}^*=c_{jt}/\gamma(i,j,k,t)$. 由上面的迭代公式可知, 其计算复杂度为 $O(n)$. 此外, Fang 等 (2006b) 也给出了中心化偏差的另一种迭代公式. 此外, 混合偏差的迭代公式类似于中心化偏差, 我们留作习题.

B. 可卷偏差迭代公式

可卷偏差也存在迭代式. 记

$$\alpha_{ij}^k\equiv|x_{ik}-x_{jk}|(1-|x_{ik}-x_{jk}|),$$

$$\delta_{ij}=\sum_{k=1}^{s}\ln\left(\frac{3}{2}-\alpha_{ij}^k\right),\quad i,j=1,\cdots,n,\ i\neq j,\ k=1,\cdots,s.$$

则可卷偏差可以表示为 $e^{\delta_{ij}}$ 的和的形式. 对于选定的列, 交换两个元素后, 计算新设计的偏差值时需要更新 $2(n-2)$ 个 δ_{ij} 值. 假设交换行向量 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_j$ 的第 k 个元素, 则对于任意异于 $\boldsymbol{x}_i$ 或 $\boldsymbol{x}_j$ 的行向量, 行对 $(\boldsymbol{x}_i, \boldsymbol{x}_t)$ 和 $(\boldsymbol{x}_j, \boldsymbol{x}_t)$ 之间的距离都会改变. 记行对 $(\boldsymbol{x}_i, \boldsymbol{x}_t)$ 和 $(\boldsymbol{x}_j, \boldsymbol{x}_t)$ 之间的距离分别为 $\tilde{\delta}_{it}$ 和 $\tilde{\delta}_{jt}$, 则

$$\tilde{\delta}_{it}=\delta_{it}+\ln\left(\frac{3}{2}-\alpha_{jt}^k\right)-\ln\left(\frac{3}{2}-\alpha_{it}^k\right),$$

$$\tilde{\delta}_{jt} = \delta_{jt} + \ln\left(\frac{3}{2} - \alpha_{it}^k\right) - \ln\left(\frac{3}{2} - \alpha_{jt}^k\right),$$

且目标函数变为

$$\nabla(\mathrm{WD}) = \frac{2}{n^2}\sum_{t \neq i,j}\left(e^{\tilde{\delta}_{it}} - e^{\delta_{it}} + e^{\tilde{\delta}_{jt}} - e^{\delta_{jt}}\right).$$

跳跃规则 搜索过程可能陷入一个局部最优解. 许多已存的随机优化算法, 如模拟退火算法 (Kirkpatrick et al., 1983) 和随机进化算法 (Jin et al., 2005), 都使用一些规则使得搜索过程跳出局部最优解. 这里存在不同的跳跃方法. 图 4.1 中的 TA 算法考虑用一系列的门限 τ_r 使之跳出局部最优解. 当 τ_r 比较大时, 很容易从局部最优解处跳出来. Fang 等 (2006b) 给出一种跳出局部最优解的新方法, 并称之为追求平衡性的启发式算法 (balance-pursuit heuristic algorithm).

此外, 2.6 节给出在 TA 算法中常用的几种偏差的下界. 在算法中, 若设计 $U_{\rm new}$ 的偏差值达到下界了, 则算法终止. 算法 4.2 给出 (Fang et al., 2003b) 中的 TA 算法的伪代码.

算法 4.2(门限接受法的伪代码)

1: 给出初值 I, J 和门限 τ_i, $i = 1, 2, \cdots, I$
2: 构造初始设计 $\boldsymbol{U}_0 \in \mathcal{U}(n; q^s)$
3: **for** $i = 1$ to I **do**
4: **for** $j = 1$ to J **do**
5: 产生 $\boldsymbol{U}_1 \in \mathcal{N}(\boldsymbol{U}_0)$ ($\boldsymbol{U}_0$ 的邻域)
6: **if** $D(\boldsymbol{U}_1)$ 达到下界, 则算法终止
7: **elseif** $D(\boldsymbol{U}_1) < D(\boldsymbol{U}_0) + \tau_i$ **then**
8: $\boldsymbol{U}_0 = \boldsymbol{U}_1$
9: **end if**
10: **end for**
11: **end for**

本质上, 给定参数 n, q 和 s 后, TA 算法是一个作用在 U-型设计集 $\mathcal{U}(n; q^s)$ 上改良的局部搜索算法. 算法从一个随机产生的 U-型设计 $\boldsymbol{U}_0$ 出发, 然后使用迭代局部搜索方法. 在每一搜索步骤中需要在当前设计的邻域中选择一个 U-型设计 $\boldsymbol{U}_1$, 并比较这两个设计的偏差值. 与简单局部搜索算法相比, 当新设计的偏差值低于或不太高于当前设计的偏差时, 都需要把当前设计更新为新设计. 这个改进可以接受一些更坏的情形, 以跳出局部最优解. 这里起决定性作用的是门限值 τ_r, 其随着迭

代次数的增加而逐渐减少至 0, 即刚开始时, 即使 $\boldsymbol{U}_1$ 的偏差值比 $\boldsymbol{U}_0$ 的大, 也要接受 $\boldsymbol{U}_1$ 为新设计, 越到后面, 趋向于只接受不大于当前设计偏差值的新设计.

轨迹图　可用于判断迭代过程的行为, 其给出目标函数在每一步迭代的取值. 图 4.2 给出某个寻找均匀设计的 TA 算法的轨迹图, 其中横轴代表迭代次数, 纵轴代表在每步迭代时的当前设计 $\boldsymbol{U}_c$ 的偏差值，例如, 混合偏差值. 由图 4.2 可知, 混合偏差值在初始阶段像一个随机游走, 然后经过多次迭代后向最优目标快速地收敛. 图 4.3 是同一问题在不同参数 I, J 和 τ_i 下的轨迹图. 读者可以从中发现其不同的迭代行为.

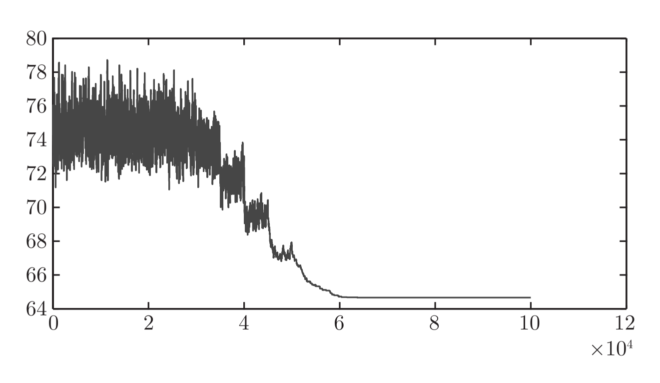


图 4.2　TA 算法轨迹图

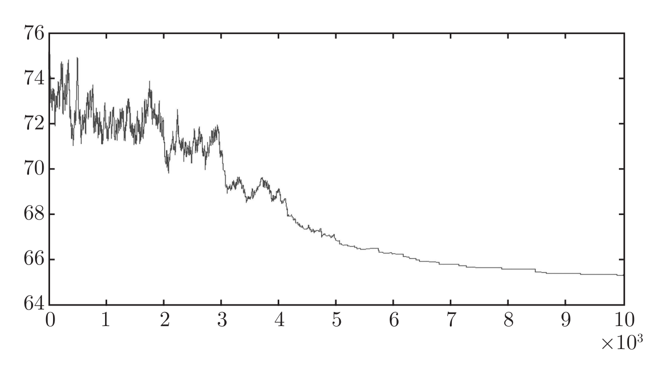


图 4.3　不同参数下的 TA 算法的轨迹图

考虑历史最优　当 $\boldsymbol{U}_0$ 为一个具有较小偏差的设计时, TA 算法可能获得一个比 $\boldsymbol{U}_0$ 更差的设计 $\boldsymbol{U}_{\rm opt}$. 这种情形会经常发生, 因为当门限值不是太小时, TA 算法鼓励当前设计 $\boldsymbol{U}_c$ 跳出局部最优解. 为了克服这个困难, Fang 等 (2016) 考虑一种称为*考虑历史最优*的新结构, 其允许 TA 算法中的当前设计 $\boldsymbol{U}_c$ 回到迭代过程中的历史最优点. 例如, 我们可以在每当门限值变化时, 嵌入一个判断命令. 当目标函数为混合偏差时, 每当 TA 算法中门限值变化时, 比较 ${\rm MD}(\boldsymbol{U}_c)$ 和 ${\rm MD}(\boldsymbol{U}_h)$ 的大

小, 其中 $\boldsymbol{U}_h$ 为 TA 算法中的历史最优解, 若 $\mathrm{MD}(\boldsymbol{U}_c) > \mathrm{MD}(\boldsymbol{U}_h)$, 则令 $\boldsymbol{U}_c = \boldsymbol{U}_h$ 并在改变后的门限值下再进行迭代. 若存在多个过程最优解, 则从中随机选取一个即可.

多阶段 TA 算法 初始设计 $\boldsymbol{U}_0$ 影响输出设计的质量. 一种选择方法是随机选择 m 个初始设计并运行 m 次 TA 算法, 得到 m 个输出设计, 从中选择一个最佳的输出设计作为最终的设计. 另一种选择方法是选取一个低偏差设计作为初始设计. 基于这种想法, 我们在不同的阶段利用 TA 算法. 首先, 利用 TA 算法获得一个输出设计并记为 $\boldsymbol{U}_1$; 然后把 $\boldsymbol{U}_1$ 作为初始设计应用 TA 算法得到输出设计 $\boldsymbol{U}_2$; 再把 $\boldsymbol{U}_2$ 作初始设计应用 TA 算法得到输出设计 $\boldsymbol{U}_3$; 类似地可以得到一系列的输出设计 $\boldsymbol{U}_k$. 当 $\mathrm{MD}(\boldsymbol{U}_k) - \mathrm{MD}(\boldsymbol{U}_{k-1}) = 0$ 时, 停止算法. 确定门限的 α 值在不同阶段分别记为 $[\alpha_1, \cdots, \alpha_k]$, 则为了保证这种策略的收敛性, 可以让这些 α 值逐渐减少.

例 4.1 我们希望在设计空间 $\mathcal{U}(27; 3^{13})$ 中得到一个均匀设计 $U_{27}(3^{13})$. 在设计空间中随机选取一个初始设计 $\boldsymbol{U}_0$ 且其 $\mathrm{MD}(\boldsymbol{U}_0) = 75.61$. 考虑用 TA 算法优化 $\boldsymbol{U}_0$. 在 TA 算法中做两方面的调整, 即考虑多阶段 TA 算法且考虑历史最优的方法. 参数设为 $I = 20$, $J = 5000$, $[\alpha_1, \cdots, \alpha_5] = [0.15\ 0.016\ 0.01\ 0.002\ 0.0005]$. 在每个阶段的 TA 算法中让 I 和 J 保持不变. 门限序列如 (4.6) 式选取. 每个阶段得到的输出设计分别为 $[\boldsymbol{U}_1^a \cdots \boldsymbol{U}_5^a]$. 最优的输出设计 $\boldsymbol{U}_{\mathrm{opt}}^a = \boldsymbol{U}_5^a$ 的混合偏差值 $\mathrm{MD}(\boldsymbol{U}_{\mathrm{opt}}^a) = 64.1888$. 该结果优于阶段 2 的输出设计, 其 $\mathrm{MD}(\boldsymbol{U}_2) = 64.3689$. 图 4.4 给出了多阶段 TA 算法的轨迹图. 从图中可知, 在经典的 TA 算法中加以考虑历史最优和多阶段方法这两个修改起到很好的效果. 另一方面, 在阶段 2 和阶段 3, 复原过程最优可以很好地消除迭代过程的不稳定性. 虽然 $\boldsymbol{U}_c$ 可能比初始的低偏差设计更差, 但是考虑历史最优的做法可以退回到过程最优处后再迭代. 这种方法可以使得 TA 算法也要考虑已获得的局部最优解, 同时这种算法可以保证在每个阶段的设计具有较好的均匀性. 在阶段 1, 由于 $\boldsymbol{U}_0$ 的混合偏差值比较大, TA 算法可以很有效地减少偏差, 而在阶段 4 和阶段 5, $\boldsymbol{U}_c$ 基本上不变化了.

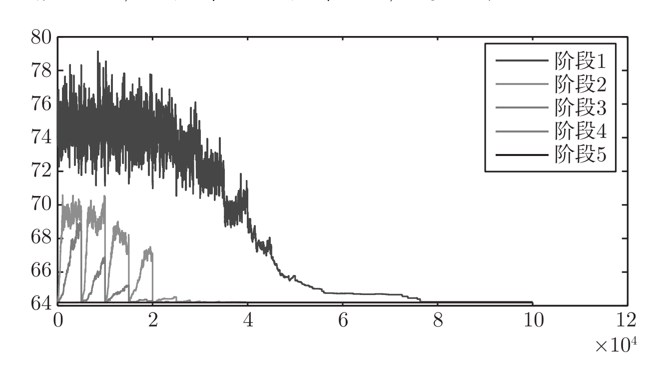


图 4.4 例 4.1 中多阶段 TA 算法在不同阶段的混合偏差的轨迹图 (文后附彩图)

4.3　用二次型方法构造均匀设计

给定偏差后, 我们可以从偏差表达式的性质出发来搜索均匀设计. 本节中, 我们将证明由 (2.36) 式定义的再生核而得到的偏差可以表示为二次型的形式, 同时把搜索均匀设计的问题转变为一个二次 0-1 整数规划问题, 并应用整数规划的求解方法来解决该问题.

4.3.1　偏差的二次型形式

在 2.3 节和 2.5 节定义的中心化偏差、可卷偏差、混合偏差和 Lee 偏差都是由 (2.36) 式的再生核定义的. 这些偏差可以表示为下面的统一形式:

$$C+\frac{1}{n^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\prod_{k=1}^{m}f(d_{ik},d_{jk},q_k)-\frac{2}{n}\sum_{i=1}^{N}\prod_{k=1}^{m}g(d_{ik},q_k),$$

其中 C 在不同偏差下是不同的常数，函数 $f(\cdot,\cdot,\cdot)$ 和 $g(\cdot,\cdot)$ 在不同的偏差下也有不同的表达式.

记 $\mathcal{V}_{q_i}=\{1,2,\cdots,q_i\}$ 且 $N=q_1\cdots q_s$, 其笛卡儿积为 $\mathcal{V}^s=\mathcal{V}_{q_1}\times\cdots\times\mathcal{V}_{q_s}$. 对于任意 U-型设计 $\mathcal{P}\in\mathcal{U}(n;q_1\times\cdots\times q_s)$, 记 $n(i_1,\cdots,i_s)$ 为点 $(i_1,\cdots,i_s)$ 在 $\mathcal{P}$ 中出现的次数. 则设计 $\mathcal{P}$ 由下面的长度为 N 的列向量唯一确定

$$\boldsymbol{y}_{\mathcal{P}}=(n(i_1,\cdots,i_s))_{(i_1,\cdots,i_s)\in\mathcal{V}^s}\,, \tag{4.7}$$

其中 $\mathcal{V}^s$ 中所有的点 $(i_1,\cdots,i_s)$ 按照**字典序**排列. $\boldsymbol{y}_{\mathcal{P}}$ 也被称为**频率向量**(Fang and Ma, 2001b; Zhou et al., 2012). 例如, 当 $q=2$ 和 $s=3$ 时, 频率向量的元素分别为 $n(1,1,1)$, $n(1,1,2)$, $n(1,2,1)$, $n(1,2,2)$, $n(2,1,1)$, $n(2,1,2)$, $n(2,2,1)$ 和 $n(2,2,2)$. 特别地, 一个完全因子设计 $U(N;q_1\times\cdots\times q_s)$ 相应的频率向量为 $\boldsymbol{y}_{\mathcal{P}}=\mathbf{1}_N$, 其中 $\mathbf{1}_N$ 为元素都是 1 的 N 维列向量. 因此

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{j=1}^{N}\prod_{k=1}^{m}f(d_{ik},d_{jk},q_k)\\=&\sum_{(i_1,\cdots,i_s)\in\mathcal{V}^s}\ \sum_{(j_1,\cdots,j_s)\in\mathcal{V}^s}n(i_1,\cdots,i_s)\,n(j_1,\cdots,j_s)\prod_{k=1}^{m}f(i_k,j_k,q_k)\end{aligned}$$

且

$$\sum_{i=1}^{N}\prod_{k=1}^{m}g(d_{ik},q_k)=\sum_{(i_1,\cdots,i_s)\in\mathcal{V}^s}n(i_1,\cdots,i_s)\prod_{k=1}^{m}g(i_k,q_k),$$

则平方偏差 $\mathrm{CD}^2(\mathcal{P})$, $\mathrm{WD}^2(\mathcal{P})$ 和 $\mathrm{MD}^2(\mathcal{P})$ 可以重写为 $\boldsymbol{y}_{\mathcal{P}}$ 的二次型.

引理 4.1 对于任一设计 $\mathcal{P}\in\mathcal{D}(n;q_1\times\cdots\times q_s)$, 有

$$\mathrm{WD}^2(\mathcal{P})=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}_{\mathcal{P}}, \tag{4.8}$$

$$\mathrm{CD}^2(\mathcal{P})=\left(\frac{13}{12}\right)^s-\frac{2}{n}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{y}_{\mathcal{P}}+\frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{y}_{\mathcal{P}}, \tag{4.9}$$

$$\mathrm{MD}^2(\mathcal{P})=\left(\frac{19}{12}\right)^s-\frac{2}{n}\boldsymbol{m}^{\mathrm{T}}\boldsymbol{y}+\frac{1}{n^2}\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y}_{\mathcal{P}}, \tag{4.10}$$

其中 $\boldsymbol{W}=\boldsymbol{W}_1\otimes\cdots\otimes\boldsymbol{W}_s$, $\boldsymbol{W}_k=(w_{ij}^k)$,

$$w_{ij}^k=1.5-|i-j|(q_k-|i-j|)/q_k^2,$$

$\boldsymbol{c}=\boldsymbol{c}_1\otimes\cdots\otimes\boldsymbol{c}_s$, $\boldsymbol{c}_k=(c_1^k,\cdots,c_{q_k}^k)^{\mathrm{T}}$, $\boldsymbol{C}=\boldsymbol{C}_1\otimes\cdots\otimes\boldsymbol{C}_s$, $\boldsymbol{C}_k=(c_{ij}^k)$,

$$c_i^k=1+|2i-1-q_k|/(4q_k)-|2i-1-q_k|^2/(8q_k^2),$$
$$c_{ij}^k=1+|2i-1-q_k|/(4q_k)+|2j-1-q_k|/(4q_k)-|i-j|/(2q_k),$$

$\boldsymbol{m}=\boldsymbol{m}_1\otimes\boldsymbol{m}_2\otimes\cdots\otimes\boldsymbol{m}_s,\boldsymbol{m}_k=(m_1^k,\cdots,m_{q_k}^k)$, $\boldsymbol{M}=\boldsymbol{M}_1\otimes\boldsymbol{M}_2\otimes\cdots\otimes\boldsymbol{M}_s$, $\boldsymbol{M}_k=(m_{ij}^k)$,

$$m_i^k=\frac{5}{3}-\frac{1}{4}\left|\frac{2i-1-q_k}{2q_k}\right|-\frac{1}{4}\left|\frac{2i-1-q_k}{2q_k}\right|^2,$$
$$m_{ij}^k=\frac{15}{8}-\frac{1}{4}\left|\frac{2i-1-q_k}{2q_k}\right|-\frac{1}{4}\left|\frac{2j-1-q_k}{2q_k}\right|-\frac{3}{4}\left|\frac{i-j}{q_k}\right|+\frac{1}{2}\left|\frac{i-j}{q_k}\right|^2,$$

$i,j=1,\cdots,q_k$, $k=1,\cdots,s$, 且 $\otimes$ 表示 Kronecker 乘积.

需要指出的是, Fang 和 Ma(2001b) 首先给出 $\mathrm{WD}^2(\mathcal{P})$ 的二次型形式, Fang 和 Qin(2003) 给出 $\mathrm{CD}^2(\mathcal{P})$ 的二次型形式, Chen 等 (2015c) 给出 $\mathrm{MD}^2(\mathcal{P})$ 的二次型形式.

4.3.2 补设计理论

一个完全因子设计 $U(N;q_1\times\cdots\times q_s)$, $N=q_1\cdots q_s$, 可以分为两个子设计 $\mathcal{P}$ 和 $\overline{\mathcal{P}}$, 其分别有 n 行和 $N-n$ 行. 这两个设计互称为补设计, 即称 $\overline{\mathcal{P}}$ 为 $\mathcal{P}$ 的补设计, 反之亦然. 记 $Z_t=\{0,1,2,\cdots,t\}$ 和 $Z_t^N=Z_t\times\cdots\times Z_t$. 则 $\boldsymbol{y}_{\mathcal{P}}\in Z_1^N$ 意味着 $\boldsymbol{y}_{\mathcal{P}}$ 的每个元素等于 0 或者 1, 且 $\boldsymbol{y}_{\mathcal{P}}=\boldsymbol{1}_N-\boldsymbol{y}_{\overline{\mathcal{P}}}$. 我们需要考虑 $\mathcal{P}$ 的偏差和 $\overline{\mathcal{P}}$ 的偏差之间的关系. 在本小节中, 我们考虑在中心化偏差、可卷偏差和混合偏差意义

下原设计和其补设计偏差之间的关系. 这些理论结果将有助于通过补设计来构造相应的大设计矩阵.

基于引理 4.1, Jiang 和 Ai(2014) 与 Chen 等 (2015c) 给出设计 $\mathcal{P}$ 的偏差与其补设计 $\overline{\mathcal{P}}$ 的偏差之间的关系. 这些结果汇总如下.

定理 4.2 对于任意设计 $\mathcal{P} \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$, $\mathcal{P}$ 的偏差与其补设计 $\overline{\mathcal{P}}$ 的偏差之间有下面的关系:

$$\begin{aligned}\mathrm{WD}^2(\mathcal{P}) = &-\left(\frac{4}{3}\right)^s + \frac{2n-N}{n^2}\prod_{k=1}^{s}\left(\frac{4q_k}{3}+\frac{1}{6q_k}\right)\\ &+\frac{(N-n)^2}{n^2}\left[\mathrm{WD}^2(\overline{\mathcal{P}}) + \left(\frac{4}{3}\right)^s\right],\end{aligned} \tag{4.11}$$

$$\begin{aligned}\mathrm{CD}^2(\mathcal{P}) = &\left(\frac{13}{12}\right)^s + \frac{1}{n^2}\left(\mathbf{1}_N^{\mathrm{T}}\boldsymbol{C}\mathbf{1}_N - 2n\boldsymbol{c}^{\mathrm{T}}\mathbf{1}_N\right)\\ &+\frac{1}{n^2}\left[-2\boldsymbol{y}_{\overline{\mathcal{P}}}^{\mathrm{T}}(\boldsymbol{C}\mathbf{1}_N - N\boldsymbol{c}) + (N-n)^2\left(\mathrm{CD}^2(\overline{\mathcal{P}}) - \left(\frac{13}{12}\right)^s\right)\right],\end{aligned} \tag{4.12}$$

$$\begin{aligned}\mathrm{MD}^2(\mathcal{P}) = &\left(\frac{19}{12}\right)^s + \frac{1}{n^2}\left(\mathbf{1}_N^{\mathrm{T}}\boldsymbol{M}\mathbf{1}_N - 2n\boldsymbol{m}^{\mathrm{T}}\mathbf{1}_N\right)\\ &+\frac{1}{n^2}\left(-2\boldsymbol{y}_{\overline{\mathcal{P}}}^{\mathrm{T}}(\boldsymbol{M}\mathbf{1}_N - N\boldsymbol{m}) + (N-n)^2\left(\mathrm{MD}^2(\overline{\mathcal{P}}) - \left(\frac{19}{12}\right)^s\right)\right),\end{aligned} \tag{4.13}$$

其中 $\boldsymbol{C}$, $\boldsymbol{c}$, $\boldsymbol{M}$ 和 $\boldsymbol{m}$ 的定义见引理 4.1.

证明 仅证明混合偏差的结果, 其他两个偏差的证明类似. 由于 $\boldsymbol{y}_{\mathcal{P}} = \mathbf{1}_N - \boldsymbol{y}_{\overline{\mathcal{P}}}$, 从 (4.10) 有

$$\begin{aligned}\mathrm{MD}^2(\mathcal{P}) &= \left(\frac{19}{12}\right)^s - \frac{2}{n}\boldsymbol{m}^{\mathrm{T}}(\mathbf{1}_N - \boldsymbol{y}_{\overline{\mathcal{P}}}) + \frac{1}{n^2}\left(\mathbf{1}_N - \boldsymbol{y}_{\overline{\mathcal{P}}}\right)^{\mathrm{T}}\boldsymbol{M}\left(\mathbf{1}_N - \boldsymbol{y}_{\overline{\mathcal{P}}}\right)\\ &= \left(\frac{19}{12}\right)^s - \frac{2}{n}\boldsymbol{m}^{\mathrm{T}}\mathbf{1}_N + \frac{2}{n}\boldsymbol{m}^{\mathrm{T}}\boldsymbol{y}_{\overline{\mathcal{P}}}\\ &\quad + \frac{1}{n^2}\left(\mathbf{1}_N^{\mathrm{T}}\boldsymbol{M}\mathbf{1}_N - 2\boldsymbol{y}_{\overline{\mathcal{P}}}^{\mathrm{T}}\boldsymbol{M}\mathbf{1}_N + \boldsymbol{y}_{\overline{\mathcal{P}}}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y}_{\overline{\mathcal{P}}}\right),\end{aligned} \tag{4.14}$$

$$\mathrm{MD}^2(\overline{\mathcal{P}}) = \left(\frac{19}{12}\right)^s - \frac{2}{N-n}\boldsymbol{m}^{\mathrm{T}}\boldsymbol{y}_{\overline{\mathcal{P}}} + \frac{1}{(N-n)^2}\boldsymbol{y}_{\overline{\mathcal{P}}}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y}_{\overline{\mathcal{P}}}. \tag{4.15}$$

结合 (4.14) 和 (4.15), 易得 (4.13) 的结果. 定理得证.

定理 4.2 中的设计 $\mathcal{P}$ 不一定是 U-型设计. 根据定理 4.2 中的关系式 (4.11)–(4.13), 易得通过补设计而得到均匀设计 $\mathcal{P}$ 的方式.

定理 4.3 对于任意设计 $\mathcal{P} \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$.

(1) 在可卷偏差意义下, 设计 $\mathcal{P}$ 是一个均匀设计当且仅当其补设计 $\overline{\mathcal{P}}$ 是 $\mathcal{D}(N-n; q_1 \times \cdots \times q_s)$ 中的均匀设计.

(2) 若对于 $\mathcal{D}(n; q_1 \times \cdots \times q_s)$ 中的任意设计, $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{1}_N - N\boldsymbol{c})$ 是一个常数, 则在中心化偏差意义下, $\mathcal{P}$ 是一个均匀设计当且仅当其补设计 $\overline{\mathcal{P}}$ 是 $\mathcal{D}(N-n; q_1 \times \cdots \times q_s)$ 中的均匀设计.

(3) 若对于 $\mathcal{D}(n; q_1 \times \cdots \times q_s)$ 中的任意设计, $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{M}\boldsymbol{1}_N - N\boldsymbol{m})$ 是一个常数, 则在混合偏差意义下, $\mathcal{P}$ 是一个均匀设计当且仅当其补设计 $\overline{\mathcal{P}}$ 是 $\mathcal{D}(N-n; q_1 \times \cdots \times q_s)$ 中的均匀设计.

定理 4.3 的证明是显然的, 故省略. 与定理 4.3(1) 中关于可卷偏差的结果相比, 定理 4.3(2) 和 (3) 中需要各增加一个约束条件, 即 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{1}_N - N\boldsymbol{c})$ 或 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{M}\boldsymbol{1}_N - N\boldsymbol{m})$ 是一个常数. 进一步, 可以发现在一些情形下, 这些条件是可以满足的, 见下面的两个推论.

推论 4.4 在中心化偏差意义下, 若下面的任一条件满足

(1) 水平 $q_1, \cdots, q_s$ 都是奇数,

(2) $s = 2$,

(3) $q_1 = \cdots = q_{s-1} = 2$,

则 $\mathcal{P} \in \mathcal{D}(n; q_1 \times \cdots \times q_s)$ 是一个均匀设计当且仅当其补设计 $\overline{\mathcal{P}}$ 是 $\mathcal{D}(N-n; q_1 \times \cdots \times q_s)$ 中的均匀设计.

证明 需要证明 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{1}_N - N\boldsymbol{c})$ 在这些情形下是一个常数. 易证

$$\boldsymbol{C}_k \boldsymbol{1}_{q_k} - q_k \boldsymbol{c}_k = \frac{1}{8q_k} \boldsymbol{1}_{q_k} I_{\{q_k \text{ 为偶数}\}}, \quad k = 1, \cdots, s,$$

其中 $I_{\{\cdot\}}$ 是一个示性函数.

(1) 当 $q_1, \cdots, q_m$ 都是奇数时, $\boldsymbol{C}_k \boldsymbol{1}_{q_k} = q_k \boldsymbol{c}_k$, $k = 1, \cdots, s$. 则有 $\boldsymbol{C}\boldsymbol{1}_N = N\boldsymbol{c}$ 且 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{1}_N - N\boldsymbol{c}) = 0$ 为常数.

(2) 当 $s = 2$ 时, 有

$$\begin{aligned}
&\boldsymbol{C}\boldsymbol{1}_N - N\boldsymbol{c} \\
=&\left(q_1\boldsymbol{c}_1 + \frac{1}{8q_1}\boldsymbol{1}_{q_1} I_{\{q_1 \text{ 为偶数}\}}\right) \otimes \left(q_2\boldsymbol{c}_2 + \frac{1}{8q_2}\boldsymbol{1}_{q_2} I_{\{q_2 \text{ 为偶数}\}}\right) - N\boldsymbol{c} \\
=&\frac{q_2}{8q_1}\boldsymbol{1}_{q_1} \otimes \boldsymbol{c}_2 I_{\{q_1 \text{ 为偶数}\}} + \frac{q_1}{8q_2}\boldsymbol{c}_1 \otimes \boldsymbol{1}_{q_2} I_{\{q_2 \text{ 为偶数}\}} \\
&+ \frac{1}{64q_1q_2}\boldsymbol{1}_N I_{\{q_1 \text{和 } q_2 \text{同为偶数}\}}.
\end{aligned}$$

由于 $\boldsymbol{y}_D^{\mathrm{T}}\mathbf{1}_N = n$, $\boldsymbol{y}_D^{\mathrm{T}}(\mathbf{1}_{q_1}\otimes \boldsymbol{c}_2)$ 和 $\boldsymbol{y}_D^{\mathrm{T}}(\boldsymbol{c}_1\otimes \mathbf{1}_{q_2})$ 对于 $\mathcal{D}(n;q_1,q_2)$ 中的任意设计都是常数, 则 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\mathbf{1}_N - N\boldsymbol{c})$ 是一个常数.

(3) 当 $q_k=2$ 时, $\boldsymbol{c}_k = \dfrac{35}{32}\mathbf{1}_2$, $k=1,\cdots,s-1$, 有 $\boldsymbol{c} = \left(\dfrac{35}{32}\right)^{s-1}\mathbf{1}_{2^{s-1}}\otimes \boldsymbol{c}_s$. 此外, $\boldsymbol{C}_k\mathbf{1}_2 = \dfrac{9}{4}\mathbf{1}_2$, $k=1,\cdots,s-1$ 且 $\boldsymbol{C}_s\mathbf{1}_{q_s} - q_s\boldsymbol{c}_s = \dfrac{1}{8q_s}\mathbf{1}_{q_s} I_{\{q_s\text{ 为偶数}\}}$. 因此

$$\begin{aligned}\boldsymbol{C}\mathbf{1}_N - N\boldsymbol{c} &= \left(\left(\frac{9}{4}\right)^{s-1}\mathbf{1}_{2^{s-1}}\right)\otimes(\boldsymbol{C}_s\mathbf{1}_{q_s}) - N\boldsymbol{c}\\ &= \left(q_s\left(\frac{9}{4}\right)^{s-1} - N\left(\frac{35}{32}\right)^{s-1}\right)1_{2^{s-1}}\otimes \mathbf{c}_s + \left(\frac{9}{4}\right)^{s-1}\frac{1}{8q_s}1_N\, I_{\{q_s\text{ 为偶数}\}}.\end{aligned}$$

对于 $\mathcal{D}(n,2^{s-1}q_s)$ 中的任意设计, 有 $\boldsymbol{y}_D^{\mathrm{T}}\mathbf{1}_N = n$ 且 $\boldsymbol{y}_D^{\mathrm{T}}(\mathbf{1}_{2^{s-1}}\otimes \boldsymbol{c}_s)$, 则 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\mathbf{1}_N - N\boldsymbol{c})$ 为常数. 结论得证.

下面考虑混合偏差. 对于设计 $\mathcal{P}\in\mathcal{D}(n;2^s)$, 由定理 4.3 中 $\boldsymbol{M}$ 和 $\boldsymbol{m}$ 的定义可知

$$\boldsymbol{M}\mathbf{1}_N - N\boldsymbol{m} = \left[\left(\frac{13}{4}\right)^s - \left(\frac{305}{96}\right)^s\right]\mathbf{1}_N,$$

则 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{M}\mathbf{1}_N - N\boldsymbol{m}) = \left[\left(\dfrac{13}{4}\right)^s - \left(\dfrac{305}{96}\right)^s\right]\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}\mathbf{1}_N = (N-n)\left[\left(\dfrac{13}{4}\right)^s - \left(\dfrac{305}{96}\right)^s\right]$ 为一个常数. 因此, 我们有下面的结论.

推论 4.5　*当 $n<N=2^s$ 且均匀性测度为混合偏差时, 设计 $\mathcal{P}\in\mathcal{D}(n;2^s)$ 是一个均匀设计当且仅当其补设计 $\overline{\mathcal{P}}$ 是 $\mathcal{D}(N-n;2^s)$ 上的均匀设计.*

进一步地, 当设计的水平数大于 2, 也可以通过补设计理论获得均匀设计或近似均匀设计. 因此, 可以通过试验点数小的均匀设计出发, 再应用补设计理论得到试验次数更大的均匀设计. 从而可以大大降低搜索试验次数很大的均匀设计的计算复杂度. 需要指出的是, 在定理 4.3 中, $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{C}\mathbf{1}_N - N\boldsymbol{c})$ 或 $\boldsymbol{y}_{\mathcal{P}}^{\mathrm{T}}(\boldsymbol{M}\mathbf{1}_N - N\boldsymbol{m})$ 是常数的条件是一个充分条件. 即使该条件不满足, 补设计理论也是有用的. 下面的例子说明这种方法是构造大均匀设计的有效方法.

例 4.2　*设 $n=6, q=3$ 和 $s=3$. 考虑构造均匀设计 $U_6(3^3)$ 和 $U_{21}(3^3)$. 从 3^3 完全因子设计的 27 个试验点中选 6 个没有重复的试验点的构成试验次数为 6 的设计, 即其设计空间为 $\mathcal{D}(6;3^3)$ 中没有重复点的子集. 在所有 $\begin{pmatrix}27\\6\end{pmatrix}=296010$ 个*

可能的设计中, 下面的设计

$$\mathcal{P} = \begin{pmatrix} 1 & 1 & 2 & 2 & 3 & 3 \\ 1 & 2 & 2 & 3 & 1 & 3 \\ 1 & 3 & 2 & 1 & 2 & 3 \end{pmatrix}^{\mathrm{T}}$$

是一个在混合偏差意义下的均匀设计, 其混合偏差值为 0.1219. 因此, 其补设计为 3^3 完全因子设计中剔除上述 6 个点而得到的设计 $\overline{\mathcal{P}}$, 其混合偏差值为 0.1097. 在 3^3 完全因子设计中选 21 个没有重复的设计点而构成设计, 可以验证, 在所有可能的 296010 种设计中, 补设计 $\overline{\mathcal{P}}$ 的混合偏差值最小, 这也意味着, $\overline{\mathcal{P}}$ 也是没有重复点的 $\mathcal{D}(21;3^3)$ 中的均匀设计.

当设计的试验次数 n 很接近 N 时, 补设计理论是构造均匀设计 $\mathcal{P}$ 的有效方法. 具体构造算法如下所示.

步骤 1 给定参数 $n, q_1, \cdots, q_s$, 通过某些算法构造试验次数为 $N-n$ 的均匀设计 $\overline{\mathcal{P}}$.

步骤 2 通过 $\overline{\mathcal{P}}$ 得到补设计 $\mathcal{P}$.

例如, 当 $N=5^6$ 且 $n=N-10$ 时, 容易通过某些算法构造一个 (近似) 均匀设计 $U_{10}(5^6)$, 如采用 TA 算法等. 从而可以通过补设计理论得到 (近似) 均匀设计 $U_{N-10}(5^6)$.

4.3.3 最优频率向量

根据 4.3.1 节中不同偏差的二次型形式, 搜索均匀设计的问题可以转变为求解一个二次整数规划问题. 在过去几十年中, 包括凸优化和二次整数规划等优化理论得到了蓬勃发展, 见 (Boyd and Vandenberghe, 2004). 在本小节中, 我们应用一些求解二次整数规划问题的有效算法来构造均匀设计.

首先, 给出二次规划、凸二次规划、半正定规划和整数规划等概念. 一个优化问题可以表示如下

$$\begin{aligned} \min \quad & f_0(\boldsymbol{y}) \\ \text{s.t.} \quad & f_i(\boldsymbol{y}) \leqslant 0, \quad i=1,\cdots,t, \\ & h_i(\boldsymbol{y}) = 0, \quad i=1,\cdots,p. \end{aligned} \tag{4.16}$$

我们希望寻找一个优化向量 $\boldsymbol{y} \in R^m$ 使其在满足 (4.16) 的约束条件下最小化目标函数 $f_0(\boldsymbol{y})$. 函数 $f_0: R^m \to R$ 也被称为代价函数. 函数 $f_i: R^m \to R$ 和 $h_i: R^m \to R$ 被分别称为不等式约束函数和等式约束函数. 当 $t=p=0$ 时, 称问题 (4.16) 为无约束问题.

当 $f_0, \cdots, f_t$ 为凸函数时, $h_i(\boldsymbol{y}) = \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y} - \boldsymbol{b}_i, i=1,\cdots,p$, 则问题 (4.16) 变为一个凸优化问题. 若目标函数 f_0 是一个二次函数, 且约束函数是线性时, 该凸优化问

题也被称为**二次规划**(quadratic programming, QP) 问题. 一个二次规划问题可以表示为下面的形式:

$$\begin{aligned}\min \quad & f_0(\boldsymbol{y}) = \frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{y} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{y}, \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{y} \leqslant \boldsymbol{h}, \\ & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, \end{aligned} \tag{4.17}$$

其中 $\boldsymbol{Q} \in R^{m\times m}$ 是一个对称实矩阵, $\boldsymbol{c} \in R^m, \boldsymbol{h} \in R^t, \boldsymbol{b} \in R^p, \boldsymbol{G} \in R^{t\times m}, \boldsymbol{A} \in R^{p\times m}$, 且 $\boldsymbol{G}\boldsymbol{y} \leqslant \boldsymbol{h}$ 意味着向量 $\boldsymbol{G}\boldsymbol{y}$ 的每个元素都不大于向量 $\boldsymbol{h}$ 的相应元素. 若 $\boldsymbol{Q}$ 为对称半正定矩阵, 则称 (4.17) 式的二次规划问题为**半正定二次规划**问题. 为了解决 (4.17) 的 QP 问题, 文献中提出了很多方法, 例如, 子梯度法 (Shor, 1985)、束方法 (Hiriart-Urruty and Lemaréchal, 1993)、内点法 (Boyd and Vandenberghe, 2004)、椭球法 (Shor, 1991) 等. 若矩阵 $\boldsymbol{Q}$ 是正定的, 则椭球法可以在多项式时间内解决该问题 (Kozlov et al., 1979). 若 $\boldsymbol{Q}$ 是不定矩阵, 则 QP 问题变成一个 NP-难问题, 即使 $\boldsymbol{Q}$ 只有一个负的特征值, 该问题也是 NP-难问题 (Pardalos and Vavasis, 1991).

在许多应用中, 向量 $\boldsymbol{y}$ 的每个元素都取自于离散集 $\{0,1\}$, 二次规划问题 (4.17) 的约束变为 $\boldsymbol{y} = (y_1, \cdots, y_m) \in \{0,1\}^m$, 即 $y_i \in \{0,1\}$. 考虑下面特殊的二次规划问题:

$$\begin{aligned}\min \quad & f_0(\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{y}, \\ \text{s.t.} \quad & \boldsymbol{y} \in \{0,1\}^m. \end{aligned} \tag{4.18}$$

称 (4.18) 式为**无约束二次0-1规划问题** (unconstrained quadratic zero-one programming problem). 在很多情形下, 需要增加一个约束, 即使得 $\boldsymbol{y}$ 的元素之和等于某个常数 n,

$$\begin{aligned}\min \quad & f_0(\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{y}, \\ \text{s.t.} \quad & \boldsymbol{1}^{\mathrm{T}}\boldsymbol{y} = n, \quad \boldsymbol{y} \in \{0,1\}^m, \end{aligned} \tag{4.19}$$

其中 $\boldsymbol{1} = (1, \cdots, 1)^{\mathrm{T}}$. Iasemidis 等 (2001) 证明二次 0-1 规划问题 (4.19) 可以退化为问题 (4.18). 由于问题 (4.18) 的约束并不是凸的, 则该问题也不是一个凸优化问题, 即二次 0-1 规划问题不是一个凸优化问题, 故求解该问题是一个 NP- 难问题. 我们需要用一些优化算法来求解.

给定参数 $q_1, \cdots, q_s$ 和 n 后, 考虑构造相应的均匀设计. 基于表达式 (4.8), 在可卷偏差意义下构造均匀设计 $U_n(q_1 \times \cdots \times q_s)$ 的问题可以表示为下面的优化问题:

$$\begin{cases}\min \quad f_0(\boldsymbol{y}) = -\left(\dfrac{4}{3}\right)^s + \dfrac{1}{n^2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}, \\ \text{s.t.} \quad \boldsymbol{1}_N^{\mathrm{T}}\boldsymbol{y} = n, \ \boldsymbol{y} \in Z_+^N, \end{cases} \tag{4.20}$$

其中 $\boldsymbol{W}$ 的定义见引理 4.1, $Z_+^N = Z_+ \times \cdots \times Z_+$, $Z_+ = \{0, 1, 2, \cdots\}$, $N = q_1 \cdots q_s$, 且 $\boldsymbol{y} = (y_1, \cdots, y_N)^{\mathrm{T}} \in Z_+^N$ 意味着 $y_i \in Z_+$. 实际上, 由约束 $\mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} = n$ 和 $\boldsymbol{y} \in Z_+^N$, 可以把取值范围由 Z_+^N 缩小至 Z_n^N, 其中 $Z_n = \{0, 1, 2, \cdots, n\}$. 因此, 问题 (4.20) 等价于下面的优化问题:

$$\begin{cases} \min & f_1(\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}, \\ \text{s.t.} & \mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} = n, \quad \boldsymbol{y} \in Z_n^N. \end{cases} \tag{4.21}$$

问题 (4.21) 是一个整数规划问题, 其并不是一个凸优化问题, 因此也是一个 NP-难问题. (Boyd and Vandenberghe, 2004) 说明凸优化问题可以在多项式时间内求解. 若去掉约束 $\boldsymbol{y} \in Z_n^N$, 得到下面的凸优化问题:

$$\begin{cases} \min & f_2(\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}, \\ \text{s.t.} & \mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} = n. \end{cases} \tag{4.22}$$

问题 (4.22) 是一个特殊的凸二次规划问题, 即半正定规划问题. 因此, 可以用凸优化的理论来求解问题 (4.22). 特别地, 有下面的结论.

定理 4.6 问题 (4.22) 的解为

$$\boldsymbol{y}^* = \frac{n}{N}\mathbf{1}_N, \tag{4.23}$$

其中 $N = q_1 \cdots q_s$.

证明 易知, 问题 (4.22) 的拉格朗日乘子为

$$L(\boldsymbol{y}, \lambda) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} + \lambda(\mathbf{1}^{\mathrm{T}}\boldsymbol{y} - n) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} + \lambda\mathbf{1}^{\mathrm{T}}\boldsymbol{y} - \lambda n, \tag{4.24}$$

其中 $\lambda \in R$ 为拉格朗日乘子中的系数. 从而, 拉格朗日对偶函数可由下式得到

$$g(\lambda) = \inf_{\boldsymbol{y}} L(\boldsymbol{y}, \lambda) = \inf_{\boldsymbol{y}} \left(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} + \lambda\mathbf{1}^{\mathrm{T}}\boldsymbol{y}\right) - \lambda n. \tag{4.25}$$

对 (4.24) 式关于 $\boldsymbol{y}$ 求导并令其为 0, 则得

$$\hat{\boldsymbol{y}} = -\frac{1}{2}\lambda\boldsymbol{W}^{-1}\mathbf{1}. \tag{4.26}$$

因此

$$g(\lambda) = -\frac{1}{4}\lambda^2\mathbf{1}^{\mathrm{T}}\boldsymbol{W}^{-1}\mathbf{1} - \lambda n, \tag{4.27}$$

其给出优化问题 (4.22) 的最优解 $\boldsymbol{y}^*$ 的下界. 故关于问题 (4.22) 的拉格朗日对偶问题为

$$\max \quad g(\lambda). \tag{4.28}$$

设 d^* 和 p^* 分别为问题 (4.22) 和拉格朗日对偶问题 (4.28) 的解. 由于问题 (4.22) 是一个半正定规划问题, 其最优对偶间隔为 0, 其意味着 $d^* = p^*$ (Boyd and Vandenberghe, 2004, 226 页). 由 (4.27) 式, 问题 (4.28) 的全局最大解为

$$\lambda^* = -\frac{2n}{\mathbf{1}^{\mathrm{T}}\boldsymbol{W}^{-1}\mathbf{1}}. \tag{4.29}$$

把 (4.29) 代入 (4.26), 可得

$$\hat{\boldsymbol{y}} = -\frac{1}{2}(\boldsymbol{W}^{-1}\mathbf{1})\left(-\frac{2n}{\mathbf{1}^{\mathrm{T}}\boldsymbol{W}^{-1}\mathbf{1}}\right). \tag{4.30}$$

同时, 易得

$$\begin{cases} \boldsymbol{W}_k\mathbf{1}_{q_k} = \left(\dfrac{4q_k}{3} + \dfrac{1}{6q_k}\right)\mathbf{1}_{q_k}, & \boldsymbol{W}\mathbf{1}_N = \prod\limits_{k=1}^{s}\left(\dfrac{4q_k}{3} + \dfrac{1}{6q_k}\right)\mathbf{1}_N, \\ \boldsymbol{W}_k^{-1}\mathbf{1}_{q_k} = \left(\dfrac{4q_k}{3} + \dfrac{1}{6q_k}\right)^{-1}\mathbf{1}_{q_k}, & \boldsymbol{W}^{-1}\mathbf{1}_N = \prod\limits_{k=1}^{s}\left(\dfrac{4q_k}{3} + \dfrac{1}{6q_k}\right)^{-1}\mathbf{1}_N. \end{cases}$$

因此, 获得全局最优解 $\boldsymbol{y}^* = \dfrac{n}{N}\mathbf{1}_N$. 定理得证.

定理 4.6 对于对称和非对称设计在可卷偏差意义下的结果可以分别见 (Fang and Ma, 2001a) 与 (Zhou et al., 2012). 根据定理 4.6, 一般地, 关于 $\boldsymbol{y}^*$ 的最佳设计 $\mathcal{P}$ 并不是一个确定性设计. 不过当 $n = kN$, $k \geqslant 1$ 时, $\boldsymbol{y} = k\mathbf{1}_N$, 即完全因子设计是在可卷偏差意义下的均匀设计.

推论 4.7　当 $n = kN$, $N = q_1 \cdots q_s$ 且 k 是一个正整数时, 完全因子设计 $\mathcal{P} \in \mathcal{U}(n; q_1 \times \cdots \times q_s)$ 是在可卷偏差意义下的均匀设计, 其可卷偏差值为

$$\mathrm{WD}^2(\mathcal{P}) = \prod_{i=1}^{s}\left(\frac{4}{3} + \frac{1}{6q_i^2}\right) - \left(\frac{4}{3}\right)^s.$$

下面考虑混合偏差. 给定 n 和 $q_1, \cdots, q_s$, 基于表达式 (4.10), 构造均匀设计的问题转变为下面的最优问题

$$\begin{cases} \min \quad -\dfrac{2}{n}\boldsymbol{m}^{\mathrm{T}}\boldsymbol{y} + \dfrac{1}{n^2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y}, \\ \text{s.t.} \quad \mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} = n, \ \boldsymbol{y} \in Z_n^N, \end{cases} \tag{4.31}$$

这里, 把常数项 $\left(\dfrac{19}{12}\right)^s$ 忽略. 若把约束 $\boldsymbol{y} \in Z_n^N$ 放宽, 且让目标函数乘以 n^2, 得到下面的凸优化问题

$$\begin{cases} \min & \boldsymbol{y}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{y} - 2n\boldsymbol{m}^{\mathrm{T}}\boldsymbol{y}, \\ \text{s.t.} & \boldsymbol{1}_N^{\mathrm{T}}\boldsymbol{y} = n. \end{cases} \tag{4.32}$$

这也是一个特殊的半正定规划问题. 我们有下面的结果.

定理 4.8 给定参数 n 和 $q_1, \cdots, q_s$, 优化问题 (4.32) 的最优解为

$$\boldsymbol{y}^* = n\left(\boldsymbol{M}^{-1}\boldsymbol{m} + \frac{1 - \boldsymbol{m}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{1}_N}{\boldsymbol{1}_N^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{1}_N}\boldsymbol{M}^{-1}\boldsymbol{1}_N\right), \tag{4.33}$$

其中 $\boldsymbol{m}$ 和 $\boldsymbol{M}$ 的定义见引理 4.1.

定理 4.8 的证明见 (Chen et al., 2015c). 由定理 4.8 可知, 最优向量 $\boldsymbol{y}^*$ 的元素可能不是完全相等的. 不过, 当 $q_1 = \cdots = q_s = 2$ 时, 元素 $m_i^k = 305/192$, 且当 $i = j$ 时, $m_{ij}^k = 7/4$, 否则 $m_{ij}^k = 3/2$. 因此

$$\boldsymbol{m}_k = \left(\frac{305}{192}, \frac{305}{192}\right),$$

$$\boldsymbol{M}_k = \begin{pmatrix} \frac{7}{4} & \frac{3}{2} \\ \frac{3}{2} & \frac{7}{4} \end{pmatrix}, \quad \boldsymbol{M}_k^{-1} = \begin{pmatrix} \frac{28}{13} & -\frac{24}{13} \\ -\frac{24}{13} & \frac{28}{13} \end{pmatrix}.$$

这里记号 $m_i^k, m_{ij}^k, \boldsymbol{m}_k$ 和 $\boldsymbol{M}_k$ 如引理 4.1 所示. 此外, 根据 Kronecker 乘积的性质,

$$(\boldsymbol{B} \otimes \boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} \otimes \boldsymbol{B}^{-1}, \quad (\boldsymbol{B} \otimes \boldsymbol{B})(\boldsymbol{b} \otimes \boldsymbol{b}) = \boldsymbol{B}\boldsymbol{b} \otimes \boldsymbol{B}\boldsymbol{b},$$
$$(\boldsymbol{B} \otimes \boldsymbol{B})\boldsymbol{1}_{2k} = \boldsymbol{B}\boldsymbol{1}_k \otimes \boldsymbol{B}\boldsymbol{1}_k,$$

则 $\boldsymbol{M}^{-1}\boldsymbol{m} = \left(\frac{305}{634}\right)^s \boldsymbol{1}_N$, $\boldsymbol{M}^{-1}\boldsymbol{1}_N = \left(\frac{4}{13}\right)^s \boldsymbol{1}_N$ 且 $\boldsymbol{y}^* = \frac{n}{N}\boldsymbol{1}_N$. 因此, 有下面的结论.

推论 4.9 在设计空间 $\mathcal{D}(n; 2^s)$ 中, 在混合偏差意义下的均匀设计的频率向量为 $\boldsymbol{y}^* = \frac{n}{N}\boldsymbol{1}_N$, 其中 $N = 2^s$. 进一步地, 两水平的完全因子设计是混合偏差下的均匀设计.

Ma 等 (2003) 给出在中心化偏差意义下类似的结果, 有兴趣的读者可以参见该文献.

4.3.4 整数规划问题方法

构造均匀设计 $U(n; q_1 \times \cdots \times q_s)$可以转变为一个优化问题 (4.21), 放宽其约束条件后可以变为问题 (4.22). 定理 4.6 给出其最优解 $\boldsymbol{y}^* = \frac{n}{N}\boldsymbol{1}_N$, 其中 $N = q_1 \cdots q_s$. 然而, 当 $n < N$ 时. 相应的设计并不是一个确定性设计. 因此, 需要一些方法构造 $n < N$ 时的均匀设计. 在本小节中, 考虑用整数规划问题方法来求解均匀设计.

不失一般性, 考虑可卷偏差为均匀性度量. 由于 $n < N$, 且均匀设计要求试验点尽可能均匀地散布在试验区域内, 则约束 $\mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} = n$ 可以转变为 $\boldsymbol{y} \in \{0,1\}^N$, 其中 $\boldsymbol{y} \in \{0,1\}^N$ 意味着 $\boldsymbol{y}$ 的每个元素 $y_i \in \{0,1\}$. 则带有线性约束的二次整数规划问题 (4.21) 可以变为下面的无约束优化问题 (Iasemidis et al., 2001)

$$\begin{cases} \min & \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} + K(\boldsymbol{W})(\mathbf{1}_N^{\mathrm{T}}\boldsymbol{y} - n)^2, \\ \text{s.t.} & \boldsymbol{y} \in \{0,1\}^N, \end{cases} \tag{4.34}$$

其中 $K(\boldsymbol{W}) = 2\sum_{i=1}^N \sum_{j=1}^N |w_{ij}| + 1$, 且 $\boldsymbol{W} = (w_{ij})$ 的定义见引理 4.1. 又根据性质 $y_i^2 = y_i$, 把问题 (4.34) 重写如下:

$$\begin{cases} \min & \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}_0\boldsymbol{y} + K(\boldsymbol{W})n^2, \\ \text{s.t.} & \boldsymbol{y} \in \{0,1\}^N, \end{cases} \tag{4.35}$$

其中 $\boldsymbol{Q}_0 = \boldsymbol{W} + K(\boldsymbol{W})\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}} - 2K(\boldsymbol{W})N\boldsymbol{I}_N$ 是一个对称矩阵. 一般地, 在大多数情形下 $K(\boldsymbol{W})$ 的值很大, 且 $\boldsymbol{Q}_0$ 的元素值也较大. 为此, 可以把目标函数除以 $K(\boldsymbol{W})$ 且去掉常数部分 $K(\boldsymbol{W})n^2$, 则问题 (4.35) 可以重写为下面的无约束二次规划问题:

$$\begin{cases} \min & f(\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{y}, \\ \text{s.t.} & \boldsymbol{y} \in \{0,1\}^N, \end{cases} \tag{4.36}$$

其中 $\boldsymbol{Q} = \boldsymbol{Q}_0/K(\boldsymbol{W})$ 是一个非负定的矩阵.

易知, 无约束 0-1 二次规划问题 (4.36) 不存在解析解, 且该问题是 NP- 难问题. 为了解决这一问题, 文献中提出许多方法. 这些方法包括一些轨迹法, 如 Tabu 搜索法 (Beasley, 1998; Palubeckis, 2004)、模拟退火 (simulated annealing, SA) 算法 (Beasley, 1998; Katayama and Narihisa, 2001); 也包括基于总体的方法, 如散布搜索法 (Amini et al., 1999)、进化算法 (Merz and Freisleben, 1999; Merz and Katayama, 2004), 以及其他一些局部启发式算法 (Merz and Freisleben, 2002). 更多的细节可参见 (Gilli and Winkler, 2009). 此外, Katayama 和 Narihisa (2001) 给出一种基于 SA 的启发式算法, 其可以应用于 500 个到 2500 个变量的情形, 并与其他启发式方法作比较. 计算结果表明, 这种 SA 算法可以在较短时间内得到高质量的解. 为了解决这个问题, Merz 和 Freisleben (2002) 提出一种贪婪启发式算法和两个局部搜索算法: 1-opt 局部搜索法和 k-opt 局部搜索法.

Zhou 等(2012) 提出了一种基于 SA 的整数规划算法 (SA-based integer programming method, SA-IPM), 用于解决构造均匀设计的这一特殊的 0-1 二次整数规划问题. SA-IPM 结合了贪婪启发式算法和 1-opt 局部搜索法, 其伪代码见算法 4.3.

算法 4.3(SA-IPM)

```
1:  给出初始值 I, J, T_init, T_f, T_r
2:  产生一个初始的随机解 y_0 ∈ {0,1}^N
3:  for i = 1 : I
4:      令 y = y_0, T = T_init, ct = 0
5:      计算 y 的收益 g_k, k ∈ {1, ··· , N}
6:      while ct < J
7:          Set ct = ct + 1
8:          for t = 1 : N
9:              确定 j 使得 g_j = min_k g_k
10:             若 g_j < 0, 则令 ct = 0 且 y_j = 1 − y_j (并更新所有的 g_i)
11:             否则随机选择 k ∈ {1, ··· , N}, 以概率 e^{−g_t/T}, 令 y_k = 1 − y_k, 并
                更新所有的 g_i
12:         end
13:         令 T = T_f × T
14:     end
15:     若由 y 相应的设计达到下界, 输出 y
16:     否则, 令 y_0 = y, T_init = T_r × T_init
17: end
18: 输出 y
```

在上面的伪代码中, 参数 I, J 和 T_{init} 分别表示退火过程的外循环次数、内循环次数和初始温度. 参数 $T_f, T_r \in (0,1)$ 是两个温度降低的比率. 一般地, 可以根据一些事先的测试来设定算法 4.3 中参数初始值, 其需要在最终得到设计的质量和计算时间之间取个平衡. Zhou 等 (2012) 建议 $T_{\text{init}} = 1/q$, 其中 q 为水平数, $J = 10$ 且当 $N < 500$ 时, 比例 $T_f = 0.99, I = 10, T_r = 0.9$; 当 $N \geqslant 500$ 时, $I = 2, T_r = 0.8$.

定义当前解 $\boldsymbol{y} = (y_1, \cdots, y_N)$ 的一个邻域如下

$$\{\boldsymbol{y}_i = (y_1, \cdots, y_{i-1}, 1 - y_i, y_{i+1}, \cdots, y_N), i = 1, \cdots, N\},$$

则 $\boldsymbol{y}_i$ 和 $\boldsymbol{y}$ 之间的 Hamming 距离等于 1. 定义收益 $g_i = f(\boldsymbol{y}_i) - f(\boldsymbol{y})$, 其中 $f(\cdot)$ 是问题 (4.36) 中的目标函数. 因此, $g_i < 0$ 意味着 $\boldsymbol{y}_i$ 是一个好解, $g_i > 0$ 意味着 $\boldsymbol{y}_i$ 不是一个好解. 根据 (Merz and Freisleben, 2002), 收益 g_i 可以计算如下

$$g_i = q_{ii}(\bar{y}_i - y_i) + 2 \sum_{k=1, k \neq i}^{m} q_{ki} y_k (\bar{y}_i - y_i), \tag{4.37}$$

其中 $\bar{y}_i = 1 - y_i$, q_{ki} 是问题 (4.36) 中的矩阵 $\boldsymbol{Q}$ 的 (k,i)- 元素. 我们可以在 N 的线性时间内计算收益 g_i, 不过需要 $O(N^2)$ 时间来计算邻域中所有的收益. 不过这些收益 g_i 不需要每次迭代都重复计算. 设当前解的所有收益已计算, 可以有效地计算新的收益 g_i^n 如下:

$$g_i^n = \begin{cases} -g_i, & i = k, \\ g_i + 2q_{ik}(\bar{y}_i - y_i)(\bar{y}_k - y_k), & \text{否则}, \end{cases} \tag{4.38}$$

因此, 可以在线性时间内更新这些收益.

在算法 4.3 中, 步骤 (9) 包含一个最快速改进目标函数的局部搜索, 其异于经典的模拟退火算法. 在步骤 (10) 中可能存在不同的 j 满足 $g_j = \min_i g_i$, 我们随机地选择其中一个. 在步骤 (10) 和 (11) 中, 所有的收益 g_i 可以用 (4.38) 式更新. 此外, 若步骤 (10) 中达到下界, 则算法终止. 可卷偏差的下界见 2.6 节.

Zhou 等 (2012) 指出, 当 $n < N$ 时, 算法 4.3 可以有效地搜索 (近似) 均匀设计. 例如, 当 $N < 3200$ 时, 与许多已存的构造算法相比, 该算法有更快的计算速度并得到更低的偏差的设计. 此外, SA-IPM 算法的计算时间跟 $N = q_1, \cdots, q_s$ 密切相关, 其随着 N 的增大而指数增加. 另一方面, 给定 N 后, SA-IPM 算法的计算时间与试验次数 n 关系不大. 因此, SA-IPM 算法非常适合构造试验次数 n 很大且总的水平组合数 N 低于几千的情形. 对于其他均匀性测度, SA-IPM 算法有类似的结果.

习　　题

4.1　牛顿法是常用的优化法. 设 $f(\boldsymbol{x})$ 是定义域 $\mathcal{X}$ 上的某一函数. 我们选择一个 $\boldsymbol{x}^* \in \mathcal{X}$ 使得 $f(\boldsymbol{x}^*) = \max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$. 列出可以应用牛顿法寻优的条件. 进一步地, 若目标函数 $f(\boldsymbol{x})$ 有多个局部最优解, 如何应用牛顿法寻找全局最优解?

4.2　简单介绍下面的优化方法: 凸规划、线性规划、二次规划、整数规划、非线性规划和随机规划.

4.3　设 $\mathcal{P}$ 为下面的 8 个试验次数、4 个两水平因子的设计.

行	1	2	3	4
1	1	1	1	1
2	2	2	2	2
3	1	1	2	2
4	2	2	1	1
5	1	2	1	2
6	2	1	2	1
7	1	2	2	1
8	2	1	1	2

给出相应的频率向量 $\boldsymbol{y}_{\mathcal{P}}$, 其定义见 (4.7).

4.4　编写算法 4.2 中的 TA 算法的程序. 该程序可以自由地选择不同偏差, 如 WD，CD 或 MD, 且给出轨迹图. 应用该程序构造均匀设计 $U_9(3^4), U_{16}(4^5), U_{18}(3^7)$ 和 $U_{25}(5^6)$.

4.5　给出混合偏差 MD 的迭代式.

4.6　编写 4.3.4 小节中 SA-IPM 程序, 并应用该程序搜索近似均匀设计 $U_9(3^4)$, $U_{45}(3^4)$, $U_{20}(5^4)$,

$U_{200}(5^4), U_{12}(4^5)$ 和 $U_{600}(4^5)$.

4.7　我们要构造均匀设计 $U_n(n^s)$. 下面的前向构造法是一个自然的想法. 第一列选择为 $\boldsymbol{u}_1 = (1, 2, \cdots, n)^{\mathrm{T}}$. 第二列 $\boldsymbol{u}_2$ 是 $\{1, 2, \cdots, n\}$ 的一个置换, 其使得 $\boldsymbol{u}_2 \neq \boldsymbol{u}_1$ 且 $[\boldsymbol{u}_1, \boldsymbol{u}_2]$ 在所有的 $n! - 1$ 个置换中有最小的偏差. 第三列 $\boldsymbol{u}_3$ 也是 $\{1, 2, \cdots, n\}$ 的一个置换, 使得 $\boldsymbol{u}_3 \neq \boldsymbol{u}_1, \boldsymbol{u}_3 \neq \boldsymbol{u}_2$ 且 $[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ 在所有的 $n! - 2$ 个置换中有最小的偏差. 持续该过程以得到 s 列. 类似的想法也可以构造 $U_n(q^s)$. 试写出该构造过程的程序, 并用于构造 $U_9(9^3), U_9(3^4)$ 和 $U_8(2^7)$. 把这些结果与已有结果相比较.

第5章 建模方法

假设 $y = f(\boldsymbol{x}) = f(x_1, \cdots, x_s)$ 为一个系统的真实模型, 其中 $X_1, \cdots, X_s$ 为区域 $\mathcal{X}$ 上取值为 $(x_1, \cdots, x_s)$ 的因子, y 为响应. 对于模型未知的实体试验, $f(\cdot)$ 是未知的; 对于计算机试验, 函数 $f(\cdot)$ 是已知的但它可能没有解析表达式. 在特定的意义下, 我们需要找到一个高质量的拟模型 (metamodel) $\hat{y} = g(\boldsymbol{x})$ 去近似真实模型. 如何找到一个拟模型是一个具有挑战性的课题. 本章结合均匀设计, 介绍了许多实用建模方法, 例如, 径向基函数、多项式回归模型、样条方法、傅里叶模型、小波基和 Kriging 模型等. 更多的讨论和建模方法可参见 (Eubank, 1988; Wahba, 1990; Hastie et al., 2001; Fang et al., 2006a).

当真实模型未知时, 统计学家 G. E. P. Box 认为 "所有的模型都是错的, 但是有些是有用的". 一个建模过程就是在区域 $\mathcal{X}$ 上寻找一个可以近似真实模型的高质量拟模型的过程. 一个好的拟模型应该有较低的计算复杂度且易于研究因子与响应或者输入变量与输出变量之间的关系. Chen 等 (2006) 认为 "系统性能的数学替代模型可用于近似系统性能和设计参数之间的关系". 我们的经验也表明在模型中仅有少数因子是显著的. 现有各种各样的建模方法. 在本章中, 5.1 节介绍基函数法, 其可用于分析实体试验和计算机试验的数据. 5.2 节考虑在计算机试验数据中广受欢迎的 Kriging 模型. 5.3 节给出一个例子来说明均匀设计的用处.

5.1 基函数

很多建模方法是基于一些特殊基函数. 令

$$\{B_1(\boldsymbol{x}), B_2(\boldsymbol{x}), \cdots\}, \quad \boldsymbol{x} \in \mathcal{X}$$

为定义在试验区域 $\mathcal{X}$ 上的一组基函数. 这里, 函数 $B_i(\boldsymbol{x})(i = 1, 2, \cdots)$ 是已知的. 近似模型 g 为备选的极大模型的子模型

$$g(\boldsymbol{x}) = \beta_1 B_1(\boldsymbol{x}) + \beta_2 B_2(\boldsymbol{x}) + \cdots + \beta_m B_m(\boldsymbol{x}), \tag{5.1}$$

其中 $\beta_j(j = 1, 2, \cdots, m)$ 是待估的未知系数, m 可以是有限的也可以是无限的. 在实际应用中, 常常使用多项式基、样条基和小波基. 本节中, 我们介绍这些基函数在建模中的应用.

5.1.1 多项式回归模型

多项式回归模型被广泛用于建模中. 一维 k 阶多项式基为

$$1, x, x^2, \cdots, x^k;$$

中心化多项式基为

$$1, (x-\bar{x}), (x-\bar{x})^2, \cdots, (x-\bar{x})^k,$$

其中 $\bar{x}$ 为 x 的样本均值. 有时, 中心化多项式基比多项式基更有用. 下面重新考虑例 1.1.

例 5.1 (例 1.1 续) 试验者想要研究一次试验中强度 (y) 和化学物质用量 (x) 之间的关系. 显然, 若仅仅观测两个或三个不同的 x 值是不够的, 相似于因子设计, 这些不同的 x 值仍被称为水平. 假设水平数为 12. 在 $[0.4, 1]$ 中等距的点为 $x = 0.4250$, 0.4750, 0.5250, 0.5750, 0.6250, 0.6750, 0.7250, 0.7750, 0.8250, 0.8750, 0.9250, 0.9750. 在表 5.1 中, 每个 x 的值重复两次且相应的强度记为 $y_{k1}, y_{k2}, k = 1, \cdots, 12$.

表 5.1 单因子 UD 和响应

试验号	x	y_{k1}	y_{k2}
1	0.4250	1.2066	1.5996
2	0.4750	1.8433	0.9938
3	0.5250	1.4580	1.1150
4	0.5750	2.0226	1.9433
5	0.6250	2.1987	2.1855
6	0.6750	2.8344	3.3655
7	0.7250	3.4903	3.5215
8	0.7750	3.1198	3.6166
9	0.8250	3.7632	3.0041
10	0.8750	1.6193	2.1473
11	0.9250	0.8490	1.3500
12	0.9750	0.1136	0.8764

在图 5.1 的 y 对 x 的关系图中, 显然简单线性模型 $y = \beta_0 + \beta_1 x + \varepsilon$ 是不合适的. 有很多可供选择的近似模型, 例如, 可以考虑适当的 k 阶多项式回归模型

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k + \varepsilon, \tag{5.2}$$

或中心化多项式回归模型

$$y = \beta_0 + \beta_1(x-\bar{x}) + \beta_2(x-\bar{x})^2 + \cdots + \beta_k(x-\bar{x})^k + \varepsilon. \tag{5.3}$$

图 5.1 给出了 5 阶和 6 阶中心化多项式回归模型的拟合模型. 表 5.2 给出了后者的方差分析表. 在表 5.1 中, 因为有 12 个重复点, 所以进行了失拟检验 (对于失拟

检验, 读者可参见文献 (Myers, 1990). 如果失拟检验是不显著的, 当前的模型可以作为近似模型; 否则, 我们应该尝试其他模型. 对于该例的数据, 失拟检验是不显著的, 所以 6 阶多项式模型可以用作近似模型. 此时 σ^2 的估计可以结合 “失拟” 和 “纯误差” 的平方和: $\hat{\sigma}^2 = (0.5856 + 1.6090)/(5 + 12) = 0.1291$. 此模型检验是显著的, 这意味着模型能够很好地拟合数据. 然而, 从图 5.1 中的图形来看, 显然由于在 0.4 和 1 的邻域里拟合得不太好, 所以可以进一步地提高该近似模型.

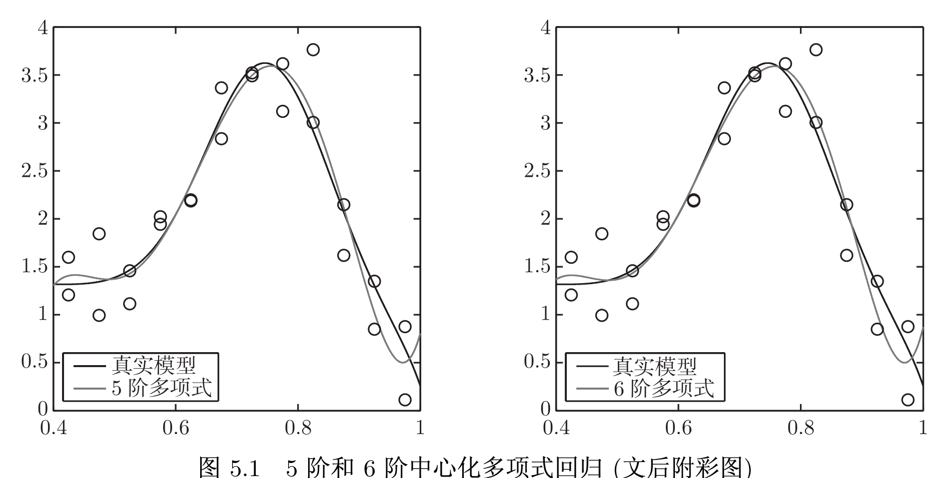


图 5.1　5 阶和 6 阶中心化多项式回归 (文后附彩图)

表 5.2　6 阶中心化多项式回归模型的方差分析表

方差来源	自由度	平方和	均方	F 值	p 值
模型	6	22.3939	3.7323	27.8322	<0.0001
失拟	5	0.5861	0.1172	0.8740	0.5266
纯误差	12	1.6090	0.1341		
总和	23	24.5889			

对于多维情况, 多项式回归模型是基于特殊的基函数 $x_1^{r_1}\cdots x_s^{r_s}$, 其中 $r_1,\cdots,r_s$ 是非负整数. 通常, 推荐低阶多项式模型, 如一阶回归模型

$$E(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_s x_s, \tag{5.4}$$

二阶回归模型

$$E(y) = \beta_0 + \sum_{i=1}^{s} \beta_i x_i + \sum_{i\leqslant j}^{s} \beta_{ij} x_i x_j \tag{5.5}$$

和中心化二阶回归模型

$$E(y) = \beta_0 + \sum_{i=1}^{s} \beta_i (x_i - \bar{x}_i) + \sum_{i\leqslant j}^{s} \beta_{ij} (x_i - \bar{x}_i)(x_j - \bar{x}_j), \tag{5.6}$$

其中 $\bar{x}_i$ 为 x_i 的样本均值.

面对高阶多项式, 我们面临两个问题: ① 随着因子数目和多项式自由度的增加, 多项式的基函数的个数极大地增加; ② 多项式基可能导致共线性问题, 即回归项之间存在较高的相关性. 在这些情况下, 推荐使用**正交多项式模型** (orthogonal polynomial model) 来克服由于多重共线性带来的困难. 单因子试验的正交回归模型是众所周知的. 对多因子的情况, 相应的模型为

$$E(y) = \alpha_0 + \sum_{i=1}^{s} \alpha_i \phi_i(x_i) + \sum_{i \leqslant j}^{s} \alpha_{ij} \phi_i(x_i) \phi_j(x_j), \tag{5.7}$$

其中 $\phi_j(u)$ 满足

$$\int_0^1 \phi_j(u)\,\mathrm{d}u = 0, \quad \int_0^1 \phi_j^2(u)\,\mathrm{d}u = 1 \quad 和 \quad \int_0^1 \phi_j(u)\phi_k(u)\,\mathrm{d}u = 0, \ 对于\ j \neq k.$$

通过位置变换, 可以利用 $[-1/2, 1/2]$ 上的 Legendre 多项式来构造 $[0, 1]$ 上单因子正交多项式. 前几个 Legendre 多项式是

$$\begin{aligned}
&\phi_0(u) = 1,\\
&\phi_1(u) = \sqrt{12}\left(u - \frac{1}{2}\right),\\
&\phi_2(u) = \sqrt{180}\left\{\left(u - \frac{1}{2}\right)^2 - \frac{1}{12}\right\},\\
&\phi_3(u) = \sqrt{2800}\left\{\left(u - \frac{1}{2}\right)^3 - \frac{3}{20}\left(u - \frac{1}{2}\right)\right\},\\
&\phi_4(u) = 210\left\{\left(u - \frac{1}{2}\right)^4 - \frac{3}{14}\left(u - \frac{1}{2}\right)^2 + \frac{3}{560}\right\},\\
&\phi_5(u) = 252\sqrt{11}\left\{\left(u - \frac{1}{2}\right)^5 - \frac{5}{18}\left(u - \frac{1}{2}\right)^3 + \frac{5}{336}\left(u - \frac{1}{2}\right)\right\},\\
&\phi_6(u) = 924\sqrt{13}\left\{\left(u - \frac{1}{2}\right)^6 - \frac{15}{44}\left(u - \frac{1}{2}\right)^4 + \frac{5}{176}\left(u - \frac{1}{2}\right)^2 - \frac{5}{14784}\right\}.
\end{aligned}$$

正交多项式连同它们的张量积

$$\Phi_{r_1,\cdots,r_s}(\boldsymbol{x}) = \prod_{j=1}^{s} \phi_{r_j}(x_j),$$

易于构造在试验区域 $\mathcal{X}$ 中的正交多项式基. 令 $B_0(\boldsymbol{x}) \equiv \Phi_{0,\cdots,0}(\boldsymbol{x}) = 1$, 且任何函数 $\Phi_{r_1,\cdots,r_s}(\boldsymbol{x})$ 的有限集合都可以用来定义此基函数.

傅里叶基函数是著名的正交基函数. 当响应 y 是因子的周期函数时, 傅里叶回归模型是很有用的. Riccomango 等 (1997) 给出关于这种基函数的详细研究. 众所周知

$$1, \cos(2\pi x), \sin(2\pi x), \cdots, \cos(2k\pi x), \sin(2k\pi x), \cdots$$

构成了一个在 $[0,1]$ 上函数空间的正交基. 实际中, 可以考虑下面的傅里叶回归模型:

$$E(y) = \beta_0 + \sum_{i=1}^{s}\sum_{j=1}^{m}\{\alpha_{ij}\cos(2j\pi x_i) + \beta_{ij}\sin(2j\pi x_i)\}.$$

当存在一些交互项时, 两因子全模型为

$$\begin{aligned} E(y) = \theta_0 &+ \sqrt{2}\sum_{r=1}^{m_1}[\sin(2\pi r x_1)\beta_{1,r} + \cos(2\pi r x_1)\gamma_{1,r}] \\ &+ \sqrt{2}\sum_{s=1}^{m_2}[\sin(2\pi s x_2)\beta_{2,s} + \cos(2\pi s x_2)\gamma_{2,s}] \\ &+ 2\sum_{r=1}^{m_1}\sum_{s=1}^{m_2}[\sin(2\pi r x_1)\sin(2\pi s x_2)\theta_{r,s} + \sin(2\pi r x_1)\cos(2\pi s x_2)\lambda_{r,s} \\ &+ \cos(2\pi r x_1)\sin(2\pi s x_2)\tau_{r,s} + \cos(2\pi r x_1)\cos(2\pi s x_2)\phi_{r,s}], \end{aligned}$$

其中 m_1 和 m_2 是傅里叶回归模型中的阶数, $\beta_{1,r}, \gamma_{1,r}$, $\beta_{2,s}, \gamma_{2,s}$, $\theta_{r,s}, \lambda_{r,s}, \tau_{r,s}, \phi_{r,s}$ 是未知参数. Riccomango 等 (1997) 指出, 在上述模型下可选择一个适当的设计使其为 D-最优的. Shi 等 (2001) 提出了 "FOUND 算法", 其中可以找到很多 D-最优的设计. 更多的讨论, 读者可以参见 (Xie et al., 2007).

5.1.2 样条基

多项式基在全试验区域 $\mathcal{X}$ 上起作用. 当函数 $g(\boldsymbol{x})$ 在 $\mathcal{X}$ 中的一些区域上变化不大, 而在 $\mathcal{X}$ 的其他区域波动较大, 多项式模型很难对这样的函数建模. 我们期望有一种特殊的基函数, 它的一些项只在 $\mathcal{X}$ 的特殊区域中出现. 例如, $I(a \leqslant x \leqslant b)$ 在 $[a,b]$ 之外的区域为零, 其中 $I(\cdot)$ 为示性函数且 $(x-a)_+$ 在区间 $(-\infty, a]$ 内为零, 其中

$$x_+ = \begin{cases} x, & x > 0, \\ 0, & \text{否则}. \end{cases}$$

对于单因子情况, 在预定的节点 $\kappa_1, \cdots, \kappa_m$ 处, *幂次样条基* (power spline basis) 具有如下的一般形式:

$$1, x, x^2, \cdots, x^p, (x-\kappa_1)_+^p, \cdots, (x-\kappa_m)_+^p. \tag{5.8}$$

广泛使用的阶数为 $p \leqslant 5$. Hastie 等 (2001) 认为 “据说对人类的眼睛来说, 三次样条是不连续节点不可见的最低阶样条”. 在实际应用中, $p=3$ 的三次样条模型是很有用的.

多因子试验的样条基函数可以通过张量积的方法来构造. 一般先对变量 x 进行标准化使得所有的变量 x 的取值范围相同, 如 [0,1]. 然后对每个变量 x 取相同的节点. 对于给定的固定节点 $\kappa_1, \cdots, \kappa_m$, 记

$$S_0(u)=1, S_1(u)=u, \cdots, S_p(u)=u^p,$$
$$S_{p+1}(u)=(u-\kappa_1)_+^p, \cdots, S_{p+m}(u)=(u-\kappa_m)_+^p.$$

则在 $\boldsymbol{x}=(x_1, \cdots, x_s)$ 上的 s 维张量积是

$$B_{r_1,\cdots,r_s}(\boldsymbol{x})=\prod_{j=0}^{s} S_{r_j}(x_j). \tag{5.9}$$

易知 $B_{0,\cdots,0}(\boldsymbol{x})=1$ 且函数 $B_{r_1,\cdots,r_s}(\boldsymbol{x})$, $0 \leqslant r_1, \cdots, r_s \leqslant m+p$, 任一个有限集构成了一个基函数. 但是, 随着维数 s 的增加, 张量积方法的基函数个数呈指数增长. Friedman (1991) 提出了所谓的*多元自适应回归样条* (multivariate adaptive regression splines, MARS). 在 MARS 中, 基函数的个数和节点的位置由数据自适应地决定. Friedman (1991) 深入地讨论了 MARS, 并给出了产生 MARS 样条基函数的算法. 读者可参见上述文献来了解其细节.

5.1.3 小波基

小波基已广泛应用于数值分析和非参数的小波压缩方法中 (参见 (Daubechies, 1992; Chui, 1992; Antoniadis and Oppenheim, 1995)). 后者经常用于试验数据的建模. 实际中, 最为广泛使用的小波压缩方法是 Donoho-Johnstone 的 VisuShrink 方法 (参见 (Donoho and Johnstone, 1994; Donoho et al., 1995)).

小波基的定义如下. 令 ϕ 和 ψ 分别为正交父函数和母小波函数. 假设 ψ 的 r 阶矩为零, ϕ 满足 $\int \phi(x)\mathrm{d}x=1$. 定义

$$\phi_{jk}(x)=2^{j/2}\phi(2^j x-k), \quad \psi_{jk}(x)=2^{j/2}\psi(2^j x-k). \tag{5.10}$$

集合 $\{\phi_{jk}(x), \psi_{jk}(x)\}$ 构成了 $L_2([0,1])$ 上的正交基, 其中 $L_2([0,1])$ 为 [0,1] 上的平方可积函数集. 这些基有相应精确的正交离散小波变换, 该变换将数据转换为小波系数域. 对给定的 $g \in L_2([0,1])$, 它可展开为小波级数

$$g(x)=\sum_{k=1}^{2^{j_0}} a_{j_0 k}\phi_{j_0 k}(x)+\sum_{j=j_0}^{\infty}\sum_{k=1}^{2^j} b_{jk}\psi_{jk}(x), \tag{5.11}$$

其中

$$a_{jk}=\int g(x)\phi_{jk}(x)\mathrm{d}x,\quad b_{jk}=\int g(x)\psi_{jk}(x)\mathrm{d}x.$$

小波变换将函数分解为不同分辨度成分. 在 (5.11) 中, a_{j_0k} 为粗糙水平系数. 它们表示函数 g 的总体结构, 然而小波系数 b_{jk} 表示函数 g 的结构随着分辨度水平 j 的增加越来越精细. 小波基把独立同分布的高斯噪声转换为独立同分布的高斯噪声且保持正态不变性. 该性质允许我们将函数域中的问题转化到小波系数域中. 关于小波基函数的系统介绍, 读者可参见 (Brown and Cai, 1997; Cai and Brown, 1998).

De Boor 和 Bon (1990) 提出的**最小插值多项式** (least interpolating polynomial) 是另一类基于基函数 (5.1) 的近似模型, 其 $B_j(\boldsymbol{x})$ 是由设计点的位置决定的多项式. 相比高阶项, 该方法倾向于低阶项. 然而, 文献中一些作者也指出该方法的不足之处, 因此我们删去了具体细节.

5.1.4　径向基函数

径向对称函数 (radial symmetric function) 有以下形式:

$$\phi(\boldsymbol{x})=\phi(x_1,\cdots,x_s)=r(||\boldsymbol{x}||),\tag{5.12}$$

其中函数 r 给定, $||\boldsymbol{x}||$ 为 $\boldsymbol{x}$ 的欧氏范数. 不同的球密度分布, 像多元正态分布、混合正态分布、在超球体 R^s 上的均匀分布、对称 Kotz- 型分布、包含多元 t- 分布和多元柯西分布的对称多元 Ⅰ 和Ⅶ型分布 (Fang et al., 1990), 给出了大量的候选模型. Hardy (1971) 提出径向基函数且 Dyn 等 (1986) 和 Powell (1987) 做了进一步的推广. 径向基函数是径向对称函数的线性组合. (5.12) 的推广形式是

$$K(||\boldsymbol{x}-\boldsymbol{x}_i||/\theta),\quad i=1,\cdots,n,\tag{5.13}$$

其中 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 是设计点, $K(\cdot)$ 是核函数, θ 是光滑参数. 推荐的近似模型形如

$$g(\boldsymbol{x})=\mu+\sum_{j=1}^{n}\beta_jK(||\boldsymbol{x}-\boldsymbol{x}_i||/\theta).\tag{5.14}$$

常用的基函数有

线性	$K(z)=z$
三次	$K(z)=z^3$
Thin-plate 样条	$K(z)=z^2\log z$
高斯	$K(z)=\exp(-z^2)$

此外, 文献中还提出许多精确插值的改进方法. 更多的细节介绍和讨论, 可以参见 Fang 等 (2006a) 的 5.6.3 小节.

5.1.5 变量选择

随着因子数的增加, 函数基中有许多函数. 当一个近似模型包含很多项时, 它将会产生:

(1) 共线性 (colinearity): 变量之间存在高度相关性;

(2) 稀疏性 (sparsity): 观测的数本数比变量个数更少;

(3) 维数祸根 (curse of dimensionality): 由于计算复杂度的问题, 许多方法不能实施.

因此, 我们不得不选择基函数的一个子集, 其需应用变量选择技术. 有很多变量选择方法, 例如, 向前选择、向后删除、逐步回归、最优子集、主成分回归、偏最小二乘回归、惩罚似然方法、贝叶斯方法. 这是一个很热门的研究领域, 在这里我们不再赘述. 读者很容易在文献中找到许多参考资料.

5.2 建模方法: Kriging 模型

在之前的章节中, 考虑以下形式的极大模型

$$y(\boldsymbol{x}_k) = \sum_{i=1}^{m} \beta_i B_i(\boldsymbol{x}_k) + \varepsilon_k, \tag{5.15}$$

其中 $\varepsilon_k (k = 1, \cdots, n)$ 独立同分布且均值为 0, 方差为 σ^2. 在很多情况下, 对误差为独立同分布的假设并不合适. 上述模型可修正如下

$$y(\boldsymbol{x}) = \sum_{i=1}^{m} \beta_i B_i(\boldsymbol{x}) + z(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{X}, \tag{5.16}$$

其中 $z(\boldsymbol{x})$ 为 $\boldsymbol{x}$ 的随机函数. 对于实体试验, 我们把 z 看成相对于测量误差的白噪声 (见定义 5.1). 然而, 在计算机试验中没有测量误差, 则 $z(\boldsymbol{x})$ 是对给定的线性模型的系统偏离.

$z(\boldsymbol{x})$ 存在许多选择, 例如, 可假设 $z(\boldsymbol{x})$ 是白噪声或高斯随机函数. 另一方面, 从计算机试验的特点考虑, 试验中没有随机误差. 我们应该有 $\hat{y}(\boldsymbol{x}_k) = y(\boldsymbol{x}_k), k = 1, \cdots, n$, 其中 $\hat{y}(\boldsymbol{x}_k)$ 是所选近似模型中 $y(\boldsymbol{x}_k)$ 的估计. 因此, 近似模型在任意 $\boldsymbol{x} \in \mathcal{X}$ 处响应的预测值实际上是通过试验 (训练) 数据上的插值 (interpolation).

在地理统计学领域的驱动下, 所谓的Kriging 模型 (Kriging model) 或者空间相关系数模型 (spatial correlation coefficient model) 是满足上述要求的建模方法. “Kriging” 一词是优化空间预测的同义词. 它被名叫 Krige 的南非采矿工程师命名. 他首次普及关于空间预测的随机方法. 在 20 世纪 70 年代早期, 他的工作被如 Matheron (1971) 进一步发展且构成了如今因地理统计学而著名的整个研究领域的

基础, 参见 (Cressie, 1988, 1993; Goovaerts, 1997). Kriging 模型已经用于计算机试验建模. 该方向上有大量的研究, 例如, 读者可参考综述 (Sacks et al., 1989a, 1989b; Welch et al., 1992). Stein (1999) 给出了 Kriging 模型更一般的理论.

5.2.1 模型

定义 5.1 令 $y(\boldsymbol{x})$ 为区域 $\mathcal{X} \subset R^s$ 上的随机过程. 称 $y(\boldsymbol{x})$ 为一个高斯随机函数, 如果对任意的 $m > 1$ 和 $\mathcal{X}$ 上任意 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$, 向量 $(y(\boldsymbol{x}_1), \cdots, y(\boldsymbol{x}_m))$ 服从协方差函数为

$$\mathrm{Cov}(y(\boldsymbol{x}_i), y(\boldsymbol{x}_j)) = \sigma^2 R(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

的多元正态分布, 其中 $\sigma^2 > 0$ 是未知的方差, $R(\cdot,\cdot)$ 是相关系数函数. 若 $R(\boldsymbol{x}_i, \boldsymbol{x}_j)$ 仅依赖于 $\boldsymbol{x}_i - \boldsymbol{x}_j$, 称高斯随机函数 $y(\boldsymbol{x})$ 为平稳的. 在这种情况下, 相关系数函数记为 $R(\cdot)$. 对于平稳高斯函数, 如果相关系数函数满足 $R(0) = 1$ 和 $R(a) = 0, a \neq 0$, 相应的高斯随机函数被称为白噪声.

高斯随机函数由均值函数 $\boldsymbol{\mu}(\boldsymbol{x}) = E(y(\boldsymbol{x}))$ 和它们的协方差函数 $\mathrm{Cov}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathrm{Cov}(y(\boldsymbol{x}_1), y(\boldsymbol{x}_2))$ 确定.

定义 5.2 Kriging 模型可表示如下:

$$y(\boldsymbol{x}) = \sum_{i=1}^{m} \beta_i h_i(\boldsymbol{x}) + z(\boldsymbol{x}), \tag{5.17}$$

其中 $h_j(\boldsymbol{x})$ 是已知函数, β_j 是未知的待估系数, $z(\boldsymbol{x})$ 是均值为 $E(z(\boldsymbol{x})) = 0$, 协方差为

$$\mathrm{Cov}(z(\boldsymbol{x}_i), z(\boldsymbol{x}_j)) = \sigma_z^2 R(\boldsymbol{x}_i - \boldsymbol{x}_j)$$

的随机平稳函数, 其中 σ_z^2 是未知方差, R 为给定的相关系数函数. $\sum_{i=1}^{m} \beta_i h_i(\boldsymbol{x})$ 称为参数项 (parametric item). 当参数项为一个常数时, 记为 μ, 此模型退化为

$$y(\boldsymbol{x}) = \mu + z(\boldsymbol{x}), \tag{5.18}$$

称其为平凡 Kriging 模型 (ordinary Kriging model), 它广泛应用于建模.

在文献中, 广泛使用幂指数相关系数函数

$$R(\boldsymbol{x}_i, \boldsymbol{x}_k) = \prod_{j=1}^{s} R_j(x_{ij}, x_{kj}) = \prod_{j=1}^{s} \exp(-\theta_j |x_{ij} - x_{kj}|^{\gamma_j}), \tag{5.19}$$

其中 $\theta_j \geqslant 0$ 是未知参数, $0 \leqslant \gamma_j \leqslant 2$ 是给定的. 令 $d_j = x_{ij} - x_{kj}$. 函数 $R_j(d_j)$ 有更多选择如下:

(1) 指数: $\exp(-\theta_j |d_j|)$;

(2) 高斯: $\exp(-\theta_j d_j^2)$;

(3) LIN: $\max(0, 1-\theta_j|d_j|)$;

(4) 样条: $1-3\xi_j^2+2\xi_j^3$, $\xi_j=\min(1,\theta_j|d_j|)$;

(5) Matérn 族:

$$\frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{2\sqrt{\nu}|d_j|}{\theta_j}\right)^{\nu}K_{\nu}\left(\frac{2\sqrt{\nu}|d_j|}{\theta_j}\right), \tag{5.20}$$

其中 θ_j 是未知参数, $\nu>0$, $K_\nu(\cdot)$ 是修正的 ν 阶 Bessel 函数, 它的定义可在维基百科中找到. 令 $\boldsymbol{\beta}=(\beta_1,\ \beta_2,\cdots,\ \beta_m)^{\mathrm{T}}$, 协方差参数为 σ_z^2, 光滑参数向量为 $\boldsymbol{\theta}=(\theta_1,\ \theta_2,\cdots,\ \theta_s)$. 如同所有相关系数函数一样, 这些函数值随着 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_k$ 之间的距离的增加而趋于零. 这表明样本数据点在预测点上的影响随着它们的分散距离的增加而减弱. $\boldsymbol{\theta}$ 的量级意味着该影响有多快地减弱. 对于较大的 $\boldsymbol{\theta}$ 值, 只有非常接近的数据点才有很高的相关性. 对于较小的 $\boldsymbol{\theta}$ 值, 远离的点仍影响该点的预测, 因为数据点之间仍具有很强的相关性. 我们经常选择 $\theta_1=\cdots=\theta_s=\theta$, 称其为**高斯核的带宽** (width of the Gaussian kernel). 参数 γ_j 为光滑性的度量. 响应的光滑性随着 γ_j 值的增加而增加. 此外, 需注意相关系数函数的插值性, 因为 $R(\boldsymbol{x}_i,\boldsymbol{x}_i)=1$, 预测点将精确地通过任何测量点.

5.2.2 估计

对给定的某一数据集 $\{(y_i,\boldsymbol{x}_i), i=1,\cdots,n\}$, 需要估计模型中的所有未知参数 $\boldsymbol{\beta}$, σ_z^2 和 $\boldsymbol{\theta}$. 为了估计这些未知参数, 线性无偏估计 $\hat{y}(\boldsymbol{x})=\boldsymbol{c}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y}$ 是合适的估计, 其中 $\boldsymbol{c}(\boldsymbol{x})=(c_1(\boldsymbol{x}),\cdots,c_n(\boldsymbol{x}))^{\mathrm{T}}$ 是常数向量, $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$ 是数据的响应向量, $\hat{y}$ 是拟合结果. 文献中, 倾向于使用以下线性预测.

定义 5.3 (a) 称 $y(\boldsymbol{x})$ 的预测 $\hat{y}(\boldsymbol{x})$ 为**线性预测** (linear predictor), 如果它具有形式 $\hat{y}(\boldsymbol{x})=\sum\limits_{i=1}^{n}c_i(\boldsymbol{x})y_i$.

(b) 称预测 $\hat{y}(\boldsymbol{x})$ 为**无偏预测** (unbiased predictor), 如果 $E\{\hat{y}(\boldsymbol{x})\}=E\{y(\boldsymbol{x})\}$.

(c) 称预测 $\hat{y}(\boldsymbol{x})$ 为**最佳线性无偏预测** (best linear unbiased predictor, BLUP), 如果它有**最小均方预测误差** (minimum mean squared prediction error, MSPE), 对所有的线性无偏预测有 MSPE $(\hat{y})=E\{\hat{y}(\boldsymbol{x})-y(\boldsymbol{x})\}^2$.

BLUP 通过下式给出

$$\hat{y}(\boldsymbol{x})=\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x})\hat{\boldsymbol{\beta}}+\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{\beta}}), \tag{5.21}$$

其中

$$\boldsymbol{h}(\boldsymbol{x})=(h_1(\boldsymbol{x}),\cdots,h_m(\boldsymbol{x}))^{\mathrm{T}}:m\times 1,$$

$$\begin{aligned}
&\boldsymbol{H} = (h_j(\boldsymbol{x}_i)) : n \times m, \\
&\boldsymbol{V} = (\mathrm{Cov}(z(\boldsymbol{x}_i), z(\boldsymbol{x}_j))) : n \times n, \\
&\boldsymbol{v}(\boldsymbol{x}) = (\mathrm{Cov}(z(\boldsymbol{x}), z(\boldsymbol{x}_1)), \cdots, \mathrm{Cov}(z(\boldsymbol{x}), z(\boldsymbol{x}_n)))^{\mathrm{T}} : n \times 1
\end{aligned}$$

和

$$\hat{\boldsymbol{\beta}} = [\boldsymbol{H}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{H}]^{-1}\boldsymbol{H}\boldsymbol{V}^{-1}\boldsymbol{y},$$

它是 $\boldsymbol{\beta}$ 的广义最小二乘估计. $\hat{y}(\boldsymbol{x})$ 的 MSPE 为

$$\mathrm{MSPE}(\hat{y}(\boldsymbol{x})) = \sigma_z^2 - (\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}), \boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x})) \begin{pmatrix} \boldsymbol{0} & \boldsymbol{H}^{\mathrm{T}} \\ \boldsymbol{H} & \boldsymbol{V} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{h}(\boldsymbol{x}) \\ \boldsymbol{v}(\boldsymbol{x}) \end{pmatrix}.$$

令 $\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{v}(\boldsymbol{x})/\sigma_z^2$, $\boldsymbol{R} = \boldsymbol{V}/\sigma_z^2$. 此时 MSPE 可表示为

$$\mathrm{MSPE}(\hat{y}(\boldsymbol{x})) = \sigma_z^2 \left[1 - (\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}), \boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})) \begin{pmatrix} \boldsymbol{0} & \boldsymbol{H}^{\mathrm{T}} \\ \boldsymbol{H} & \boldsymbol{R} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{h}(\boldsymbol{x}) \\ \boldsymbol{r}(\boldsymbol{x}) \end{pmatrix}\right]. \quad (5.22)$$

σ_z^2 的极大似然估计为

$$\hat{\sigma}_z^2 = \frac{1}{n}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{\beta}})^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{\beta}}).$$

注意到上述估计中包含未知参数 $\boldsymbol{\theta}$. 估计 $\boldsymbol{\theta}$ 的一种方式是极小化关于 $\boldsymbol{\theta}$ 的 MSPE $(\hat{y}(\boldsymbol{x}))$. 更多的细节讨论和例子, 读者可以参考 (Sacks et al., 1989a; Miller and Frenklach, 1983; Santner et al., 2003).

5.2.3 极大似然估计

如果考虑 Kriging 模型, 极大似然估计是估计 $\boldsymbol{\beta}$, σ_z^2 和 $\boldsymbol{\theta}$ 的一种方法. 为强调参数 $\boldsymbol{\theta}$, 记 $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{\theta})$. 在这种情况下, $\boldsymbol{y}$ 的密度为

$$(2\pi\sigma_z^2)^{-n/2}|\boldsymbol{R}(\boldsymbol{\theta})|^{-1/2}\exp\left\{-\frac{1}{2\sigma_z^2}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{\theta})^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\beta})\right\}. \quad (5.23)$$

去掉一个常数后, 数据的对数似然函数为

$$l(\boldsymbol{\beta}, \sigma_z^2, \boldsymbol{\theta}) = -\frac{n}{2}\log(\sigma_z^2) - \frac{1}{2}\log|R(\boldsymbol{\theta})| - \frac{1}{2\sigma_z^2}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{\theta})^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\beta}).$$

极大化对数似然函数得到 $(\boldsymbol{\beta}, \sigma_z^2, \boldsymbol{\theta})$ 的极大似然估计. 实际中, 同时极大化 $(\boldsymbol{\beta}, \sigma_z^2, \boldsymbol{\theta})$ 有一些技术上的困难. 经验观测表明, 基于同时极大化 $(\boldsymbol{\beta}, \sigma_z^2, \boldsymbol{\theta})$ 的预测表现几乎与分别估计 $\boldsymbol{\beta}$ 和 $(\sigma_z^2, \boldsymbol{\theta})$ 相同. 这与上述的理论分析一致. 实际中, 我们可以用以下的方式迭代估计 $\boldsymbol{\beta}$ 和 $(\sigma_z^2, \boldsymbol{\theta})$:

步骤 1 初始化 $\boldsymbol{\theta}$. 选择一个好的初始值 $\boldsymbol{\theta}=\boldsymbol{\theta}_0$, 它需要一些额外的信息. 事实上, 显然 $\boldsymbol{\beta}$ 的初始值可设为其最小二乘估计. 这等价于设置 $\boldsymbol{\theta}$ 的初始值为 $\mathbf{0}$. 可以选择 $\boldsymbol{\theta}_0=\mathbf{0}$.

步骤 2 $\boldsymbol{\beta}$ 的极大似然估计. 对于给定的 $\boldsymbol{\theta}_0$, $\boldsymbol{\beta}$ 的极大似然估计为

$$\hat{\boldsymbol{\beta}}_0=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{y}. \tag{5.24}$$

步骤 3 σ_z^2 的估计. 对于给定的 $\boldsymbol{\theta}_0$, σ_z^2 的极大似然估计为

$$\hat{\sigma}_z^2=n^{-1}(\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{\beta}}_0)^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{\theta}_0)(\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{\beta}}_0), \tag{5.25}$$

它为 σ_z^2 的一个有偏估计.

步骤 4 $\boldsymbol{\theta}$ 的估计. $\boldsymbol{\theta}$ 的极大似然估计没有显式表达式, Newton-Raphson 算法或 Fisher 得分算法可以用来搜索解. 找到一个解作为新的 $\boldsymbol{\theta}_0$ 并返回到步骤 2. 迭代步骤 2 到步骤 4 直到它收敛.

5.2.4 参数经验 Kriging

令 $\boldsymbol{e}_i$ 为第 i 个坐标为 1 其他元素均为 0 的单位列向量. 当 $\boldsymbol{x}$ 的第 i 个观测为 $\boldsymbol{x}_i$, 从 (5.21) 可得

$$\begin{aligned}\hat{y}(\boldsymbol{x}_i)&=\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}+\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{\beta})\\&=\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}+\boldsymbol{e}_i^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{\beta})=y_i,\end{aligned}$$

这意味着 Kriging 预测是对数据进行了插值. 该性质对于没有随机误差的计算机试验数据建模是合适的. 但是对于实体试验的数据是不合适的. 因此, 对于后者可采用 "参数经验 Kriging 模型" 来分析数据 (Sacks et al., 1989a), 该方法的细节可参阅 (Santner et al., 2003) 的附录 C. 令

$$y(\boldsymbol{x})=\sum_{i=1}^{m}\beta_i h_i(\boldsymbol{x})+z(\boldsymbol{x})+\epsilon(\boldsymbol{x})=\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{\beta}+z(\boldsymbol{x})+\varepsilon(\boldsymbol{x}), \tag{5.26}$$

其中 $\boldsymbol{h}(\boldsymbol{x}),\boldsymbol{\beta}$ 和 $z(\boldsymbol{x})$ 与之前有相同的含义和假设, $\varepsilon(\boldsymbol{x})$ 记为在 $\boldsymbol{x}$ 处均值为 0, 方差为 σ^2 的随机误差. 假设 $\varepsilon(\boldsymbol{x})$ 和 $z(\boldsymbol{x})$ 是不相关的, $\varepsilon(\boldsymbol{x})$ 是白噪声 (见定义 5.1). 令 $v(\boldsymbol{x})=z(\boldsymbol{x})+\varepsilon(\boldsymbol{x})$. 显然, $v(\boldsymbol{x})$ 是均值为 0, 协方差函数为

$$\boldsymbol{\Sigma}_v=\sigma_z^2\boldsymbol{R}(\boldsymbol{\theta})+\sigma^2\boldsymbol{I}=\sigma_z^2\left[\boldsymbol{R}(\boldsymbol{\theta})+\frac{\sigma^2}{\sigma_z^2}\boldsymbol{I}\right]=\sigma_z^2\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)$$

的高斯随机函数, 其中 $\alpha=\sigma^2/\sigma_z^2$, $\boldsymbol{R}(\boldsymbol{\theta})$ 是关于未知参数 $\boldsymbol{\theta}$ 的 $z(\boldsymbol{x})$ 的相关系数函数, 且

$$\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)=\boldsymbol{R}(\boldsymbol{\theta})+\frac{\sigma^2}{\sigma_z^2}\boldsymbol{I}=\boldsymbol{R}(\boldsymbol{\theta})+\alpha\boldsymbol{I},$$

其中 $\boldsymbol{I}$ 是单位矩阵. 在 (5.23) 中, 尽管此时 $\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)$ 不是一个相关系数矩阵, 然而 $\boldsymbol{\Sigma}_v=\sigma_z^2\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)$ 为协方差矩阵, 依然用 $\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)$ 替代 $\boldsymbol{R}(\boldsymbol{\theta})$. 我们仍可以应用上述估计参数 $(\boldsymbol{\beta},\sigma_z^2,\boldsymbol{\theta},\alpha)$ 的步骤, 其中 $\tilde{\boldsymbol{R}}(\boldsymbol{\theta},\alpha)$ 包含了未知参数 $\boldsymbol{\theta}$ 和 α. 则 σ^2 的极大似然估计可以通过 $\hat{\sigma}^2=\hat{\alpha}\hat{\sigma}_z^2$ 获得.

MATLAB Kriging 工具包中的 DACE 包可用来实施上述计算.

5.2.5 例子和讨论

本小节将 Kriging 方法应用到之前研究过的一些例子中. 首先, 继续利用参数经验 Kriging 对例 1.1 进行建模.

例 5.2 (例 1.1 和例 5.1 续) 在 1.1 节中, 利用多项式回归模型来拟合数据. 考虑一个简单模型

$$y(x)=\boldsymbol{\beta}_0+\boldsymbol{\beta}_1x+z(x)+\varepsilon(x), \tag{5.27}$$

其中 $\varepsilon(x)$ 是均值为 0、方差为 σ^2 的白噪声, $z(x)$ 是均值为 0、协方差函数为 $\sigma_z^2R(\theta)$ 的高斯随机函数, 且

$$R(d)=\exp(-\theta d^2).$$

参数的极大似然估计为 $\hat{\boldsymbol{\beta}}_0=-0.3411$, $\hat{\boldsymbol{\beta}}_1=-0.2599$, $\hat{\theta}=1.000$, $\hat{\alpha}=0.1768$, $\hat{\sigma}_z^2=0.6626$ 和 $\hat{\sigma}^2=\alpha\sigma_z^2=0.1171$. 对于 σ^2 的估计, 此模型与例 5.1 中方差分析模型具有一致性, 且后者的估计为 $\hat{\sigma}^2=0.1291$. 图 5.2 显示参数经验 Kriging 模型 (虚线) 有极好的拟合效果. 图 5.2 也显示了 Kriging 模型

$$y(x)=\boldsymbol{\beta}_0+\boldsymbol{\beta}_1x+z(x)$$

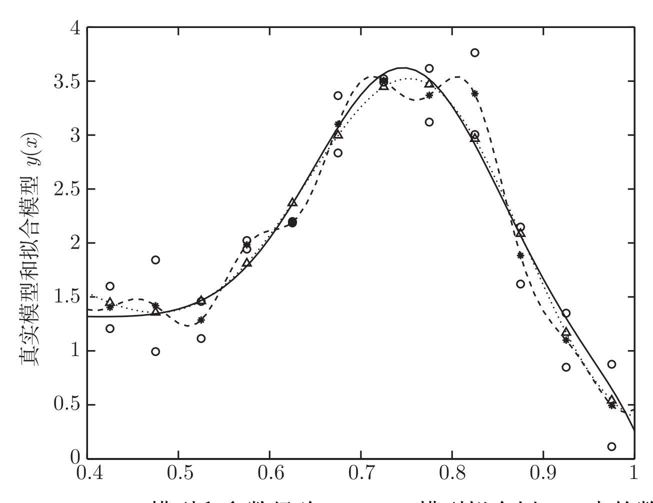


图 5.2 Kriging 模型和参数经验 Kriging 模型拟合例 1.1 中的数据

的拟合结果. 显然, 后者 (带 * 虚线) 不如经验 Kriging 模型拟合得好.

例 5.3 (例 1.13 续) 将参数经验 Kriging 模型运用到例 1.13 中. 根据例 1.13 中的讨论, 考虑以下三种模型:

$$y(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \tag{5.28}$$

$$\begin{aligned} y(\boldsymbol{x}) = &\beta_0 + \beta_1 x_2 + \beta_2 x_3 + \beta_3 x_1 x_3 + \beta_4 x_2 x_4 \\ &+ \beta_5 x_2^2 + z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \end{aligned} \tag{5.29}$$

$$\begin{aligned} y(\boldsymbol{x}) = &\beta_0 + \beta_1 (x_1 - \bar{x}_1) + \beta_2 (x_2 - \bar{x}_2) + \beta_3 (x_4 - \bar{x}_4) \\ &+ \beta_4 (x_3 - \bar{x}_3)(x_4 - \bar{x}_4) + \beta_5 (x_2 - \bar{x}_2)^2 + z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}). \end{aligned} \tag{5.30}$$

相应的参数估计为

模型 (5.28):

$$\begin{aligned} &\hat{\boldsymbol{\beta}} = (-0.0019, 0.7378, 0.3147, -0.1118, 0.3737), \\ &\hat{\boldsymbol{\theta}} = (0.8620, 0.4418, 1.5232, 0.3715), \\ &\hat{\sigma}_z^2 = 0.00024831 \quad \text{和} \quad \hat{\sigma}^2 = 0.00023053. \end{aligned}$$

模型 (5.29):

$$\begin{aligned} &\hat{\boldsymbol{\beta}} = (-0.0043, 0.9900, -0.8630, 1.0699, 0.6673, -1.2060), \\ &\hat{\boldsymbol{\theta}} = (0.6771, 0.8524, 1.000, 1.000, 1.000), \\ &\hat{\sigma}_z^2 = 0.000037869 \quad \text{和} \quad \hat{\sigma}^2 = 0.000037869. \end{aligned}$$

模型 (5.30):

$$\begin{aligned} &\hat{\boldsymbol{\beta}} = (0.0011, 0.7354, 0.3092, 0.3697, 0.00058, -0.000082), \\ &\hat{\boldsymbol{\theta}} = (0.2726, 0.2500, 1.2968, 0.2838, -0.4125), \\ &\hat{\sigma}_z^2 = 0.000043704 \quad \text{和} \quad \hat{\sigma}^2 = 0.00003675. \end{aligned}$$

表 5.3 列出了响应的拟合值以及相应的残差, 其中模型 (1)、模型 (2) 和模型 (3) 分别代表模型 (5.28)、模型 (5.29) 和模型 (5.30). 从这些结果可知:

(1) 在三个模型中, 模型 (5.28) 有较大的残差绝对值, 并且 $\hat{\sigma}_z^2$ 和 $\hat{\sigma}^2$ 的方差是最大的.

(2) 后两个模型对 $\hat{\sigma}_z^2$ 和 $\hat{\sigma}^2$ 的估计分别在同一个数量级, 但都比模型 (5.28) 中的估计小. 这些结论和例 1.13 中的结论是一致的.

(3) 在模型 (5.28) 中, 将删除 $z(\boldsymbol{x})$ 后的模型记为模型 (5.28_0). 类似地, 得到模型 (5.29_0) 和模型 (5.30_0). 相比模型 (5.28_0), 我们发现参数经验 Kriging 模型 (5.28)

有更小的估计误差的方差 $\hat{\sigma}^2$. 这表明参数经验 Kriging 模型可以更好地拟合模型. 模型 (5.29) 和 (5.30) 也有类似的结论.

(4) 模型 (5.26) 的参数项是很重要的. 许多已有的建模方法可以用来挑选参数项.

表 5.3 三种模型对例 1.13 中 y 和相应残差的拟合值

观测值	模型 (1)		模型 (2)		模型 (3)	
	$\hat{y}$	e_i	$\hat{y}$	e_i	$\hat{y}$	e_i
0.1836	0.1859	−0.0023	0.1790	0.0046	0.1810	0.0026
0.1739	0.1540	0.0199	0.1707	0.0032	0.1758	−0.0019
0.0900	0.0850	0.0050	0.0827	0.0073	0.0851	0.0049
0.1176	0.1358	−0.0182	0.1154	0.0022	0.1180	−0.0004
0.0795	0.0731	0.0064	0.0850	−0.0055	0.0699	0.0096
0.0118	0.0145	−0.0027	0.0101	0.0017	0.0168	−0.0050
0.0991	0.0873	0.0118	0.1056	−0.0065	0.0995	−0.0004
0.1319	0.1301	0.0018	0.1289	0.0030	0.1303	0.0016
0.0717	0.0725	−0.0008	0.0706	0.0011	0.0743	−0.0026
0.0109	0.0303	−0.0194	0.0131	−0.0022	0.0134	−0.0025
0.1266	0.1279	−0.0013	0.1276	−0.0010	0.1322	−0.0056
0.1424	0.1426	−0.0002	0.1504	−0.0080	0.1429	−0.0005

5.3 一个环境数据案例研究–模型选择

例 1.4 介绍了一个关于确定六种金属的浓度影响毒性的环境问题. 环保主义者认为某些水中金属元素的浓度会直接影响人体健康. 感兴趣的是研究六种金属: 镉 (Cd)、铜 (Cu)、锌 (Zn)、镍 (Ni)、铬 (Cr) 以及铅 (Pb) 的浓度和小鼠某种细胞的死亡率之间的关系. 此试验采用了均匀设计 $U_{17}(17^6)$ (见表 5.4 的左侧部分), 其中所有的金属都有 17 个水平. 对于每个设计点, 重复进行了三次试验, 并且将响应 (死亡率) 在表 5.4 中列出 (此数据被 Fang 和 Wang (1994) 采用), 其中最后一行对应的最高金属浓度的水平组合下的死亡率高于其他组. 这暗示了金属浓度可能影响死亡率. 在完成试验和收集数据之后, 研究员进行数据分析以了解死亡率与金属浓度之间的关系.

在六种金属中, 其最大浓度与最小浓度之比为 2000. 所以, 需要对数据进行标准化处理. 令 $x_1, \cdots, x_6$ 分别为 Cd, Cu, Zn, Ni, Cr 和 Pb 进行标准化后的变量. 所以, $x_1 = (\text{Cd} - 6.5624)/7.0656$, 其中 6.5624 和 7.0656 分别为因子 Cd 的样本均值和样本标准差. 类似地, 其余五个变量也进行此标准化. 此外, 对于每个设计点的

三个响应, 记 y 为这三个响应的平均值.

表 5.4 一个利用均匀设计的环境试验以及相应的死亡率

Cd	Cu	Zn	Ni	Cr	Pb	Y_1	Y_2	Y_3
0.01	0.2	0.8	5.0	14.0	16.0	19.95	17.6	18.22
0.05	2.0	10.0	0.1	8.0	12.0	22.09	22.85	22.62
0.1	10.0	0.01	12.0	2.0	8.0	31.74	32.79	32.87
0.2	18.0	1.0	0.8	0.4	4.0	39.37	40.65	37.87
0.4	0.1	12.0	18.0	0.05	1.0	31.90	31.18	33.75
0.8	1.0	0.05	4.0	18.0	0.4	31.14	30.66	31.18
1.0	8.0	2.0	0.05	12.0	0.1	39.81	39.61	40.80
2.0	16.0	14.0	10.0	5.0	0.01	42.48	41.86	43.79
4.0	0.05	0.1	0.4	1.0	18.0	24.97	24.65	25.05
5.0	0.8	4.0	16.0	0.2	14.0	50.29	51.22	50.54
8.0	5.0	16.0	2.0	0.01	10.0	60.71	60.43	59.69
10.0	14.0	0.2	0.01	16.0	5.0	67.01	71.99	67.12
12.0	0.01	5.0	8.0	10.0	2.0	32.77	30.86	33.70
14.0	0.4	18.0	0.2	4.0	0.8	29.94	28.68	30.66
16.0	4.0	0.4	14.0	0.8	0.2	67.87	69.25	67.04
18.0	12.0	8.0	1.0	0.1	0.05	55.56	55.28	56.52
20.0	20.0	20.0	20.0	20.0	20.0	79.57	79.43	78.48

Fang 和 Wang (1994) 运用逐步回归进行建模, 其模型为

$$\begin{aligned}\hat{y} = &\ 32.68 + 5.03\log(x_1) + 3.48\log(x_2) + 2.03\log(x_4) + 0.55(\log(x_2))^2\\ &-0.63(\log(x_3))^2 + 0.94(\log(x_4))^2 + 0.53\log(x_1)\log(x_2)\\ &-0.70\log(x_1)\log(x_5) + 0.92\log(x_2)\log(x_6),\end{aligned}$$

其中由于六个变量的变化范围太大, 故对这些变量取对数后再进行拟合. 我们发现逐步回归的结果十分不稳定, 即回归模型在加入或者删除变量时, 相应的 F-值十分敏感. 此现象在逐步回归建模中时常出现. 另一种方法是由 Fan 和 Li (2001) 提出的一种惩罚最小二乘方法, 叫做 SCAD. AIC 和 BIC 是这种方法的两个特例. 此方法的理论与算法实现的详细内容见 (Fan and Li, 2001).

此线性模型包含所有的线性项、二次项和二阶交互项, 加上截距项, 此模型共有 28 个待估参数. 但本试验总共仅有 17 个不同的水平组合, 不足以估计全模型. 所以, 必须要进行变量筛选. 这里用 SCAD 方法进行变量筛选. 表 5.5 中列出了待估参数的估计值 (尺度变化后的变量). 最终的回归模型包含了 12 个变量 ($t_{0.005}(38) = 2.7116$), 其中所有变量都是显著的. 从表 5.5 中可以看出, Cd 具有二

次效应, Cu 和 Ni 的效应都是线性的, 另外, Cu 与 Cr 以及 Ni 与 Cr 之间的二阶交互效应都是负的. Cr 具有负的线性效应, 并且 Cr 分别与 Cu, Zn 和 Ni 的二阶交互效应都是负的. Zn 具有二次效应. 最后, Pb 具有正的线性效应和二次效应, 并且 Pb 与 Zn 之间的二阶交互效应是正的.

表 5.5 模型参数的估计值、标准差以及 t-值

X-变量	估计值	标准差	$\|t\|$
截距	36.4539	0.5841	62.4086
Cd	14.9491	0.2944	50.7713
Cu	12.8761	0.2411	53.4060
Ni	0.9776	0.2510	3.8950
Cr	−7.2696	0.2474	29.3900
Pb	4.0646	0.2832	14.3536
Cd^2	−6.2869	0.3624	17.3480
Zn^2	2.8666	0.3274	8.7554
Pb^2	9.2251	0.4158	22.1856
Cu*Cr	−1.6788	0.3171	5.2945
Zn*Cr	−6.2955	0.3306	19.0401
Zn*Pb	11.9110	0.2672	44.5708
Ni*Cr	−11.3896	0.4303	26.4680

现在, 我们试着用参数经验 Kriging 模型对此数据进行建模. 考虑四种不同的模型: 第一种是简单的线性模型, 第二种是含有线性项的参数经验 Kriging 模型, 第三种是由 Li (2002) 推荐的模型, 第四种模型结合了第三种模型和参数经验 Kriging 模型. 具体模型如下:

$$y(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon(\boldsymbol{x}), \tag{5.31}$$

$$\begin{aligned} y(\boldsymbol{x}) = &\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \\ &+ \beta_6 x_6 + z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \end{aligned} \tag{5.32}$$

$$\begin{aligned} y(\boldsymbol{x}) = &\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_4 + \beta_4 x_5 + \beta_5 x_6 + \beta_6 x_1^2 + \beta_7 x_3^2 \\ &+ \beta_8 x_6^2 + \beta_9 x_2 x_5 + \beta_{10} x_3 x_5 + \beta_{11} x_3 x_6 + \beta_{12} x_4 x_5 + \varepsilon(\boldsymbol{x}), \end{aligned} \tag{5.33}$$

$$\begin{aligned} y(\boldsymbol{x}) = &\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_4 + \beta_4 x_5 + \beta_5 x_6 + \beta_6 x_1^2 + \beta_7 x_3^2 + \beta_8 x_6^2 \\ &+ \beta_9 x_2 x_5 + \beta_{10} x_3 x_5 + \beta_{11} x_3 x_6 + \beta_{12} x_4 x_5 + z(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \end{aligned} \tag{5.34}$$

其中 x_1—x_6 是标准化后的变量. 表 5.6 列出了模型 (5.31) 和 (5.32) 中未知参数的估计值, 而表 5.7 列出了 y 的拟合值以及相应的残差. 模型 (5.33) 和 (5.34) 的结果分别列在表 5.8 和表 5.9 中. 通过这些结果, 得到以下结论.

本试验重复进行了三次, 所以纯误差方差估计可以通过

$$\hat{\sigma}^2 = \frac{1}{2 \times 17} \sum_{i=1}^{17} \sum_{j=1}^{3} (y_{ij} - \bar{y}_i)^2 = 1.2324$$

得到. 模型 (5.33) 和 (5.34) 中 σ^2 的估计值分别是 1.1071 和 0.9140. 二者与纯误差方差的估计值 1.2324 都十分接近.

表 5.6 模型 (5.31) 和 (5.32) 的估计

X-变量	模型 (5.31)	模型 (5.32)	模型 (5.32)
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\theta}$
截距	42.8639	−0.0394	
x_1	10.7108	0.6758	0.0010
x_2	8.1125	0.3747	0.3200
x_3	−2.1168	−0.1278	4.0637
x_4	3.8514	0.2418	2.3702
x_5	−0.3443	0.0111	0.0864
x_6	0.3579	0.0562	0.9406
σ_z^2			123.2917
σ^2			1.1071

表 5.7 模型 (5.31) 和 (5.32) 中 y 的拟合值以及相应的残差

$\bar{y}$	模型 (5.31)		模型 (5.32)	
	$\hat{y}$	e_i	$\hat{y}$	e_i
18.5900	26.6164	−8.0264	18.6129	−0.0229
22.5200	23.4063	−0.8863	22.5214	−0.0014
32.4667	42.2368	−9.7701	32.4939	−0.0273
39.2967	45.0475	−5.7509	39.3332	−0.0365
32.2767	30.7434	1.5333	32.2691	0.0076
30.9933	27.4271	3.5662	30.9871	0.0062
40.0733	33.3073	6.7660	40.0573	0.0161
42.7100	46.1736	−3.4636	42.7081	0.0019
24.8900	30.9297	−6.0397	24.9044	−0.0144
50.6833	40.4781	10.2052	50.6551	0.0282
60.2767	38.4284	21.8483	60.2051	0.0715
68.7067	54.4104	14.2963	68.6476	0.0591
32.4433	44.4368	−11.9935	32.4846	−0.0412
29.7600	40.0015	−10.2415	29.7992	−0.0392
68.0533	60.0874	7.9659	68.0277	0.0256
55.7867	62.9681	−7.1814	55.8050	−0.0183
79.1600	81.9878	−2.8278	79.1751	−0.0151
MSE		9.2695		0.0316

表 5.8　模型 (5.33) 和 (5.34) 的估计

X-变量	模型 (5.33)	模型 (5.34)	
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\theta}$
截距	36.4538	−0.0006	
x_1	14.9491	0.8330	0.0064
x_2	12.8760	0.7175	0.0043
x_4	0.9776	0.0545	0.0029
x_5	−7.2695	−0.4051	0.0019
x_6	4.0644	0.2265	0.0013
x_1^2	−6.2869	−0.3292	0.0010
x_3^2	2.8666	0.1501	0.0010
x_6^2	9.2251	0.4830	0.0010
x_2x_5	−1.6788	−0.1098	0.0010
x_3x_5	−6.2954	−0.4131	0.0010
x_3x_6	11.9109	0.7777	0.0010
x_4x_5	−11.3897	−0.7473	0.0010
σ_z^2			0.7192
σ^2			0.9140

表 5.9　模型 (5.33) 和 (5.34) 中 y 的拟合值和相应的残差

$\bar{y}$	模型 (5.33)		模型 (5.34)	
	$\hat{y}$	e_i	$\hat{y}$	e_i
18.5900	18.1872	0.4028	18.1876	0.4024
22.5200	22.0938	0.4262	22.0941	0.4259
32.4667	33.0771	−0.6104	33.0767	−0.6100
39.2967	39.1439	0.1528	39.1439	0.1527
32.2767	32.1827	0.0939	32.1828	0.0938
30.9933	31.5229	−0.5296	31.5223	−0.5290
40.0733	40.0871	−0.0138	40.0872	−0.0139
42.7100	42.3616	0.3484	42.3619	0.3481
24.8900	25.1529	−0.2629	25.1528	−0.2628
50.6833	50.9327	−0.2494	50.9322	−0.2489
60.2767	60.1480	0.1287	60.1482	0.1285
68.7067	68.7445	−0.0378	68.7445	−0.0378
32.4433	32.3263	0.1170	32.3265	0.1168
29.7600	29.8031	−0.0431	29.8030	−0.0430
68.0533	67.6520	0.4013	67.6525	0.4008
55.7867	56.0953	−0.3087	56.0950	−0.3083
79.1600	79.1755	−0.0155	79.1754	−0.0154
MSE		0.3041		0.3038

在四个模型中, 参数经验模型 (5.32) 拥有最小的 MSE 值. 可以肯定的是, 当选择较好的参数项时, 参数经验模型可以使拟合的效果变好. 模型 (5.33) 和 (5.34) 的 MSE 值的数量级相同, 但都比模型 (5.32) 的 MSE 值大. 这暗示了如何选择一个好的参数项是不易的, 并且这也是一个有待解决的问题.

另外, 还有许多其他建模方法, 例如, 贝叶斯方法、包括贝叶斯高斯 Kriging 模型、神经网络、局部多项式回归等. 读者可以参考 (Fang et al., 2006a).

习　　题

5.1　假设响应 y 和因子 x 隐含如下关系

$$y = f(x) + \boldsymbol{e} = 1 - e^{-x^2} + \varepsilon, \quad \varepsilon \sim N(0, 0.1^2), \quad x \in [0, 3],$$

但试验者不知道此关系, 并且他/她想利用试验找到一个近似模型. 所以, 他/她考虑以下四种设计:

$$\begin{aligned}
D_3 &= \{0,0,0,0,1,1,1,1,2,2,2,2\},\\
D_4 &= \left\{\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{3}{4},\frac{3}{4},\frac{3}{4},\frac{5}{4},\frac{5}{4},\frac{5}{4},\frac{7}{4},\frac{7}{4},\frac{7}{4}\right\},\\
D_6 &= \left\{\frac{1}{6},\frac{1}{6},\frac{3}{6},\frac{3}{6},\frac{5}{6},\frac{5}{6},\frac{7}{6},\frac{7}{6},\frac{9}{6},\frac{9}{6},\frac{11}{6},\frac{11}{6}\right\},\\
D_{12} &= \left\{\frac{1}{12},\frac{3}{12},\frac{5}{12},\frac{7}{12},\frac{9}{12},\frac{11}{12},\frac{13}{12},\frac{15}{12},\frac{17}{12},\frac{19}{12},\frac{21}{12},\frac{23}{12}\right\}.
\end{aligned}$$

执行以下步骤:

1. 绘制函数图像

$$y = f(x) = 1 - e^{-x^2}, \quad x \in [0, 3].$$

2. 利用统计模拟为每个设计生成一个数据集.

3. 为每个数据集找到合适的回归模型以及相应的方差分析表. 绘制这些拟合模型的图像.

4. 对于每个模型, 随机产生 $N = 1000$ 个点 $x_1, \cdots, x_{1000}$ 并计算均方误差 (MSE), 其中

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2,$$

其中 $\hat{y}_i$ 是模型对 y_i 的估计值.

5. 根据绘制的图像 MSE 和 SS_E, 比较这些模型并得出你的结论.

5.2　为了比较计算机试验中不同种类的设计与建模方法, 常常考虑几种情况的研究. 模型是已知的. 选择几种设计如正交设计 (OD)、拉丁超立方抽样 (LHS) 和均匀设计 (UD) 以及一些建模方法. 然后比较所有的设计–建模组合.

假设以下模型是给定的. 请考虑三种不同的设计 OD, UD 和 LHS (其中 $n = 16, 25, 29, 64$) 和一些建模方法: 二次回归模型, (5.8) 式中生成形式为 $1, x, x^2, \cdots, x^p$, $(x-\kappa_1)_+^p, \cdots, (x-$

$\kappa_m)_+^p$ 的幂样条基, 定义 5.2 中的 Kriging 模型 $y(\boldsymbol{x}) = \sum_{i=1}^m \beta_i h_i(\boldsymbol{x}) + z(\boldsymbol{x})$ 和人工神经网络. 请比较所有的设计–建模组合.

模型 1

$$Y = \frac{\ln(x_1) \times (\sin(x_2) + 4)}{e^{x_3}} + \ln(x_1) e^{x_3}, \tag{5.35}$$

其中独立变量的取值范围分别为 $x_1 : [0.1, 10]$, $x_2 : [-\pi/2, \pi/2]$ 和 $x_3 : [0, 1]$.

模型 2

$$\begin{aligned} Y = & -\left[2 \exp\left\{-\frac{1}{2}(x_1^2 + (x_2 - 4))^2\right\}\right. \\ & \left. + \exp\left\{-\frac{1}{2}(x_1 - 4)^2 + \frac{x_2^2}{4}\right\} + \exp\left\{-\frac{1}{2}\left(\frac{(x_1+4)^2}{4} + x_2^2\right)\right\}\right], \end{aligned} \tag{5.36}$$

其中独立变量的取值范围分别为 $x_1 : [-10, 7]$, $x_2 : [-6, 7]$.

模型 3

$$\begin{aligned} Y = & 10(x_2 - x_1^2)^2 + (1 - x_1)^2 + 9(x_4 - x_3^2) + (1 - x_3)^2 \\ & + 1.01[(x_2 - 1)^2 + (x_4 - 1)^2] + 1.98(x_2 - 1)(x_4 - 1)^2, \end{aligned} \tag{5.37}$$

其中独立变量的取值范围为 $x_i : [-2, 2], i = 1, 2, 3, 4$.

模型 4

$$Y = \sum_{k=1}^{4} [100(x_{k+1} - x_k^2)^2 + (1 - x_k)^2], \tag{5.38}$$

其中独立变量的取值范围为 $x_i : [-2, 2], i = 1, \cdots, 5$.

5.3 例 1.7 是一个用来比较许多设计–建模组合不错的选择. 令 y 为机器人手臂的末端到起始部分的距离, 其可表示成 $2m$ 个变量 $\theta_j \in [0, 2\pi]$ 和 $L_j \in [0, 1]$ 的函数, 其中 $y = \sqrt{u^2 + v^2}$, 例 1.7 定义了 (u, v). 考虑三种不同的设计 OD, UD 和 LHS 以及三种近似模型: 多项式回归模型、Kriging 模型和经验 Kriging 模型. 请分别给出当 $m = 2$ 和 $m = 3$ 时可能的设计–建模组合之间的比较.

第 6 章 均匀性与其他设计准则的联系

许多试验设计关注设计的随机性、因子与每个因子水平数之间的平衡性、正交性、有效性以及稳健性, 如简单随机设计、随机区组设计、拉丁方设计、部分因子设计 (简写为 FFD)、最优设计和稳健设计. 在前面的章节中我们已经了解到均匀性 (uniformity) 在均匀设计的评价和构造中的重要地位, 本章将介绍均匀性与其他设计准则之间的紧密联系.

6.1 均匀性与同构

设 $D(n;q^s)$ 为具有 n 次试验、s 个 q 水平因子的因子设计, $L_n(q^s)$ 为正交设计, 显然它也是一类特殊的因子设计. 读者可以在 1.3.2 小节查阅因子设计的基本概念.

定义 6.1 若一个因子设计可以由另一个因子设计通过适当的行置换、列置换及水平置换得到, 那么称这两个因子设计是同构的.

由于两个同构的设计在拟合经典的 ANOVA 模型时具有相同的统计性质, 因此认为它们是等价的. 在实际应用中, 确定设计的同构性是十分重要的.

例 6.1 假设试验涉及 5 个 2 水平因子且试验员希望用 Plackett-Burman 设计 $L_{12}(2^{11})$ 来安排试验, 我们就需要从 $L_{12}(2^{11})$ 中选择一个 5 列的子设计使得该子设计在某种意义下具有最优的统计性质. 不幸的是, 由于 Plackett-Burman 设计是非正规设计, 我们不能利用最小低阶混杂准则(参见 1.4.4 小节)来选出一个子设计, 于是 Lin 和 Draper(1992) 将 $\binom{11}{5}$=462 个子设计分成两个同构类, 分别用 $L12-5.1$ 和 $L12-5.2$ 表示. 子类 $L12-5.1$ 包含 66 个设计, 而另一类 $L12-5.2$ 包含 396 个设计. 表 6.1 给出了 $L_{12}(2^5)$ 的两个有代表性的非同构子设计, 其中一个设计有一对重复行, 另一个设计没有任何重复行. 从几何角度看, 试验员更倾向于子类 $L12-5.2$, 因为它比 $L12-5.1$ 多一个自由度. 根据同构的定义, 为了判断两个 $D(n;q^s)$ 设计是否同构, 遍历搜索需要比较共 $n!(q!)^s s!$ 个设计. 例如, 欲判断两个正交 $L_{12}(2^5)$ 设计是否同构, 我们需要进行 $12!5!2^5 \approx 1.8394\times 10^{12}$ 次比较, 那么能否降低计算的复杂性?

易找到判别非同构设计的一些必要条件:

(1) 两个同构的设计 有相同的广义字长型;

(2) 两个同构的设计有相同的字母型 (letter pattern)(详情参见 (Draper and

Mitchell, 1968));

(3) 两个同构设计的任意两个不同行之间的 Hamming 距离有相同的分布.

表 6.1 两个非同构的 $L_{12}(2^{11})$ 设计

编号	$L12-5.1$					$L12-5.2$				
1	1	1	2	1	1	1	1	2	1	1
2	2	1	1	2	2	2	1	1	2	1
3	1	2	1	1	2	1	2	1	1	2
4	2	1	2	1	2	2	1	2	1	1
5	2	2	1	2	1	2	2	1	2	1
6	2	2	2	1	1	2	2	2	1	2
7	1	2	2	2	1	1	2	2	2	1
8	1	1	2	2	2	1	1	2	2	2
9	1	1	1	2	1	1	1	1	2	2
10	2	1	1	1	1	2	1	1	1	2
11	1	2	1	1	2	1	2	1	1	1
12	2	2	2	2	2	2	2	2	2	2

由定义 1.9 设计类 $L12-5.1$ 和 $L12-5.2$ 的广义字长型分别为

$$(0,0,10/9,5/9,0) \quad 和 \quad (0,0,10/9,5/9,4/9),$$

且这两类设计的分辨度均为III. 但因为它们是非同构的, 设计类 $L12-5.1$ 比 $L12-5.2$ 有着更低阶的广义混杂.

然而, 两个相同字长型的设计有可能是非同构的. Draper 和 Mitchell (1968) 给出了两个字长型完全相同的 $L_{512}(2^{12})$ 正交设计, 但它们却不是同构的.

Draper 和 Mitchell (1970) 给出了一个判定同构性更敏锐的准则, 称为"字母型准则", 并列出了 1024 行分辨度为 6 的设计. 设 a_{ij} 为正规设计 $\boldsymbol{U}$ 中包含字母 i 的长度为 j 的字的个数, $\boldsymbol{U}^a=(a_{ij})$ 为设计 $\boldsymbol{U}$ 的字母型矩阵. 他们推测两个设计 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 是同构的当且仅当 $\boldsymbol{U}_1^a=\boldsymbol{P}\boldsymbol{U}_2^a$, 其中 $\boldsymbol{P}$ 是置换矩阵. 显然, 若两个设计的字母型矩阵相同, 那么它们的字长型一定相同. Chen 和 Lin (1991) 给出了两个字母型矩阵相同的非同构设计 2^{31-15}, 于是也说明了"字母型矩阵"准则并不是设计同构的充分条件. 需要注意的是, 广义字长型和字母型矩阵都不便于计算且只能应用于较小的因子设计.

Clark 和 Dean (2001) 给出了设计同构的充要条件. 设 $\boldsymbol{H}=(d_{ij})$ 为设计 $\boldsymbol{U}$ 的 Hamming 距离矩阵, 其中 d_{ij} 是 $\boldsymbol{U}$ 的第 i 行和第 j 行之间的 Hamming 距离. 该方法的提出基于以下的事实.

引理 6.1 设 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 为两个 $D(n;q^s)$ 设计, 则 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 是同构的当且仅当存在一个 $n\times n$ 的置换矩阵 $\boldsymbol{R}$ 和 $\{1,\cdots,s\}$ 的一个置换 $\{c_1,\cdots,c_s\}$, 使得对任

意 $p=1,\cdots,s$ 有

$$\boldsymbol{H}_{\boldsymbol{U}_1}^{\{1,\cdots,p\}}=\boldsymbol{R}\boldsymbol{H}_{\boldsymbol{U}_2}^{\{c_1,\cdots,c_p\}}\boldsymbol{R}^{\mathrm{T}},$$

其中 $\boldsymbol{H}_{\boldsymbol{U}}^{\{c_1,\cdots,c_p\}}$ 是由设计 $\boldsymbol{U}$ 的 $\{c_1,\cdots,c_p\}$ 列组成的子设计的 Hamming 距离矩阵.

该巧妙的方法在水平置换下是不变的, 但其复杂度仍然使得计算难以实施.

容易发现两个同构的两水平 U-型设计有相同的 CD 值; 两个同构的两水平或三水平 U-型设计有相同的 WD 值; 两个同构的 q 水平 U-型设计有相同的离散偏差. 回顾上面的两个设计类 $L12-5.1$ 和 $L12-5.2$, 它们的 CD 值分别为 0.166541 和 0.166527, 显然这两类设计是非同构的. 但是有时两个设计即使有相同的偏差值, 它们也是非同构的, 在这种情况下子设计的投影偏差起到很大的作用.

定义 6.2 对于一个给定的 $D(n;q^s)$ 设计 $\boldsymbol{U}$ 和整数 k $(1\leqslant k\leqslant s)$, 共有 $\binom{s}{k}$ 个 $D(n;q^k)$ 子设计. 设 D 是一种度量均匀性的偏差, 从而这些子设计的 D 值可以定义一个分布, 表示为 $F_D^k(\boldsymbol{U})$. 则称该分布为设计 $\boldsymbol{U}$ 的 k 边 D 值分布 (k-marginal D-value distribution).

符号 $F_{\mathrm{WD}}^k(\boldsymbol{U})$, $F_{\mathrm{CD}}^k(\boldsymbol{U})$ 和 $F_{\mathrm{MD}}^k(\boldsymbol{U})$ 分别表示当偏差 D 取 WD, CD 或 MD 时的 k 边 D 值分布. Ma 等 (2001) 提出了下面针对同构性的均匀性准则 (uniformity criterion for isomorphism, UCI).

引理 6.2 两个 $D(n;2^s)$ 设计 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 同构的必要条件为

(a) 它们有相同的 CD_2 值;

(b) 对任意的 $1\leqslant k<s$ 它们有相同的分布 $F_{\mathrm{CD}}^k(\boldsymbol{U}_1)=F_{\mathrm{CD}}^k(\boldsymbol{U}_2)$.

基于该引理他们提出了下面的算法来识别非同构 $D(n;2^s)$ 设计, 称为 NIU 算法 (NIU algorithm). 设 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 是两个这样的设计.

算法 6.1 (NIU 算法)

步骤 1 比较 $\mathrm{CD}(\boldsymbol{U}_1)$ 和 $\mathrm{CD}(\boldsymbol{U}_2)$, 若 $\mathrm{CD}(\boldsymbol{U}_1)\neq\mathrm{CD}(\boldsymbol{U}_2)$, 则可知 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 是非同构的并终止程序, 否则转入步骤 2.

步骤 2 对 $k=1,s-1,2,s-2,\cdots,\lfloor s/2\rfloor,s-\lfloor s/2\rfloor$, 其中 $\lfloor x\rfloor$ 表示小于 x 的最大整数, 比较 $F_{\mathrm{CD}}^k(\boldsymbol{U}_1)$ 和 $F_{\mathrm{CD}}^k(\boldsymbol{U}_2)$. 若 $F_{\mathrm{CD}}^k(\boldsymbol{U}_1)\neq F_{\mathrm{CD}}^k(\boldsymbol{U}_2)$, 则可知 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 是非同构的并终止程序, 否则用下一个 k 值重复该步骤.

例如, 我们可以把该算法应用到 (Chen and Lin, 1991) 中研究过的两个 $L_{32768}(2^{31})$ 设计. 程序显示：两个设计有相同的 CD 值 4.279; 对 $k=1,30,2,29$ 所有的 k 维子设计有相同的 CD 值分布; 但是 $F_{\mathrm{CD}}^{28}(\boldsymbol{U}_1)\neq F_{\mathrm{CD}}^{28}(\boldsymbol{U}_2)$, 于是只需运行算法的几个步骤就可以证明两个设计不是同构的. 该例说明在识别非同构的设计时 NIU 算法是十分有效的.

以上的思想和算法可以扩展到识别更高水平的因子设计 (参见 (Ma et al., 2001))

和研究设计的投影性质 (参见 (Lin and Draper, 1992)).

接下来考虑如何判别水平数 q 大于 2 的非同构设计 $D(n;q^s)$ 的问题. 对一个 $D(n;q^s)$ 设计 $\boldsymbol{U}$, 设

$$E_j(\boldsymbol{U}) = \frac{1}{n}\text{card}\{(i,k)|d^H(\boldsymbol{u}_i,\boldsymbol{u}_k)=j\}, \tag{6.1}$$

其中 $\text{card}(\mathcal{A})$ 是集合 $\mathcal{A}$ 的基数, $d^H(\boldsymbol{u}_i,\boldsymbol{u}_k)$ 是两行 $\boldsymbol{u}_i$ 和 $\boldsymbol{u}_k$ 之间的 Hamming 距离, 称向量 $(E_0(\boldsymbol{U}),\cdots,E_s(\boldsymbol{U}))$ 为 $\boldsymbol{U}$ 的**距离分布** (distance distribution). 称

$$B_a(\boldsymbol{U}) = \sum_{i=1}^{s} E_i(\boldsymbol{U})a^i$$

为 $\boldsymbol{U}$ 的**距离计数** (参见 (Roman, 1992, 226 页)), 其中 a 是一个正数. 对于两水平设计 D, Ma 等 (2001) 给出了距离计数和均匀性之间的联系, 如下式

$$\text{CD}^2(D) = \frac{1}{n}\left(\frac{5}{4}\right)^s B_{4/5}(D) - 2\left(\frac{35}{32}\right)^s + \left(\frac{13}{12}\right)^s.$$

事实上, 对于两水平设计, UCI(参见引理 6.2) 与 a 取 4/5 时的测度 $B_a(\boldsymbol{U})$ 是等价的. 很自然地, 该测度可以应用到高水平因子设计. 给定 k $(1\leqslant k\leqslant s)$, 用 $F_{B_{a,k}}(\boldsymbol{U})$ 表示所有 k 维投影子设计的 B_a 值的分布, 于是可以得到高水平设计下的 NIU 算法. 由于这里参数 a 是一个提前确定的值, 为了简便起见, 我们在记号中省去 a.

算法 6.2 (高水平设计下的 NIU 算法)

步骤 1 比较 $B(\boldsymbol{U}_1)$ 和 $B(\boldsymbol{U}_2)$, 若 $B(\boldsymbol{U}_1)\neq B_2(\boldsymbol{U}_2)$, 则可知两个设计 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 不是同构的并终止程序, 否则进入步骤 2.

步骤 2 对 $k=1,s-1,2,s-2,\cdots,\lfloor s/2\rfloor,s-\lfloor s/2\rfloor$, 比较 $F_{B_k}(\boldsymbol{U}_1)$ 和 $F_{B_k}(\boldsymbol{U}_2)$, 若 $F_{B_k}(\boldsymbol{U}_1)\neq F_{B_k}(\boldsymbol{U}_2)$, 则可知 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 不是同构的并终止程序, 否则用下一个 k 值重复该步骤.

举个简单的例子, 考虑表 6.2 中列出的来自其他文献的四个 $L_{18}(3^7)$ 设计. 例如, 当 $a=4/5$ 时四个设计有相同的距离计数 6.685248, 但所有 6 维投影设计的距离计数的分布是不同的, 如表 6.3 所示, 因此可知设计 (a), (c) 和 (d) 不是同构的. 需要注意到完全比较后可以发现设计 (a) 和 (b) 实际上是同构的. 根据以往的经验, 我们更倾向于设定参数 a 为无理数.

我们知道, 两水平正交设计和 Hammard 矩阵有着紧密的联系. 若一个 Hammard 矩阵可以由另一个通过行置换、列置换或/和水平置换得到, 则称两个 Hammard 矩阵是等价的. Hammard 矩阵在构造试验设计、代码理论和其他领域中都有着举足轻重的地位. 但随着 n 的增加, 依定义进行完全搜索来判断两个 Hammard 矩阵是否等价是一个 NP- 难问题. 对于阶数 1, 2, 4, 8 和 12, 其 Hammard 矩阵有唯一的一个等价类.

表 6.2 四个 $L_{18}(3^7)$ 设计

编号	(a)	(b)	(c)	(d)
1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	1 2 1 1 3 2 1	3 1 2 1 3 1 2
2	1 2 2 2 2 2 2	1 1 2 3 2 3 1	1 3 1 1 2 1 2	1 1 2 3 1 3 1
3	1 3 3 3 3 3 3	1 2 1 3 3 2 2	1 1 2 3 2 2 3	2 2 2 3 2 1 3
4	2 1 1 2 2 3 3	1 2 3 2 1 3 3	1 3 2 3 1 3 1	1 2 3 1 1 2 2
5	2 2 2 3 3 1 1	1 3 2 2 3 1 3	1 1 3 2 3 3 2	3 1 3 1 2 3 3
6	2 3 3 1 1 2 2	1 3 3 1 2 2 2	1 2 3 2 1 1 3	1 2 1 1 2 1 1
7	3 1 2 1 3 2 3	2 1 1 2 3 3 2	2 1 1 3 1 1 2	1 3 2 2 2 3 2
8	3 2 3 2 1 3 1	2 1 3 3 1 2 3	2 2 1 3 3 3 3	2 3 1 1 3 3 1
9	3 3 1 3 2 1 2	2 2 2 2 2 2 1	2 2 2 2 2 2 2	2 1 3 2 2 2 1
10	1 1 3 3 2 2 1	2 2 3 1 3 1 1	2 3 2 2 3 1 1	3 3 1 3 2 2 2
11	1 2 1 1 3 3 2	2 3 1 3 2 1 3	2 1 3 1 1 2 1	3 2 2 2 3 2 1
12	1 3 2 2 1 1 3	2 3 2 1 1 3 2	2 3 3 1 2 3 3	1 3 3 2 3 1 3
13	2 1 2 3 1 3 2	3 1 2 1 3 2 3	3 1 1 2 2 3 1	2 2 3 3 3 3 2
14	2 2 3 1 2 1 3	3 1 3 2 2 1 2	3 3 1 2 1 2 3	3 3 3 3 1 1 1
15	2 3 1 2 3 2 1	3 2 1 1 2 3 3	3 1 2 1 3 1 3	1 1 1 3 3 2 3
16	3 1 3 2 3 1 2	3 2 2 3 1 1 2	3 2 2 1 1 3 2	2 1 1 2 1 1 2
17	3 2 1 3 1 2 3	3 3 1 2 1 2 1	3 2 3 3 2 1 1	3 2 1 2 1 3 3
18	3 3 2 1 2 3 1	3 3 3 3 3 3 1	3 3 3 3 3 2 2	2 3 2 1 1 2 3

表 6.3 四个 $L_{18}(3^7)$ 设计的 F_{B_6} 值

$B_{(a)}$	频数	$B_{(b)}$	频数	$B_{(c)}$	频数	$B_{(d)}$	频数
7.6683	1	7.6683	1	7.6719	1	7.6737	3
7.6765	6	7.6765	6	7.6747	2	7.6765	4
				7.6765	4		

Fang 和 Ge (2004) 提出了一个很有效的算法, 能够根据对称的 Hamming 距离序列轻易地辨别出不等价的 Hammard 矩阵, 并发现对于阶数 36 的 Hammard 矩阵有至少 382 个等价类.

6.2 均匀性与正交性

正交性在试验设计中发挥着重要作用, 并且有着多种不同的含义. 本节中, "正交性" 是在正交表意义下而言的. 例如, 强度为 2 的正交表就要求任一因子的诸水平之间和任两个因子的水平组合之间都有一个好的平衡性, 换句话说, 它要求一维和二维的投影均匀性. 而均匀设计通常就是关注一维投影和 s 维整体的均匀性的. 所以正如本节和 6.4 节中所提到的, 这两类设计之间应该有一些联系.

通过数值搜索 (Fang and Winker, 1998) 与 (Fang et al., 2000b) 中所提到的设计, 可以发现很多行数较少的 CD 测度下均匀设计也是正交设计, 比如 $U_4(2^3), U_8(2^7)$, $U_{12}(2^{11}), U_{16}(2^{15})$, $U_9(3^4), U_{12}(3 \times 2^3)$, $U_{16}(4^5), U_{16}(4 \times 2^{12})$, $U_{18}(2 \times 3^7)$ 和 $U_{25}(5^6)$.

该事实表明很多正交设计同时在 CD 意义下也是均匀设计, 它们可以通过数值搜索的方法来发现这些设计. 由此他们推测*任何正交设计也是某种偏差意义下的均匀设计*. Ma 等 (2003) 在 CD 意义下对该猜想进行了研究. 考虑一个好格子点集 $\mathcal{P}=\{\boldsymbol{x}_k, k=1,\cdots,n\}\in\mathcal{D}(n;q^s)$. 用 $n(i_1,\cdots,i_s)$ 表示 $\mathcal{P}$ 中水平组合 $(i_1,\cdots,i_s)$ 的个数. 公式 (4.9) 中用 $\boldsymbol{y}_{\mathcal{P}}$ 的二次型来表示 $\mathrm{CD}(\mathcal{P})$ 的平方. 当 $q_1=\cdots=q_s$ 时, 公式 (4.9) 可以简化为

$$[\mathrm{CD}(\mathcal{P})]^2=\left(\frac{13}{12}\right)^s-\frac{2}{n}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{y}+\frac{1}{n^2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{y}, \tag{6.2}$$

其中 $\boldsymbol{y}(\mathcal{P})$ (或者简写为 $\boldsymbol{y}$) 是各元素按字典序排列的 q^s 维向量, $\boldsymbol{c}_0=(c_1,\cdots,c_q)^{\mathrm{T}}$, $\boldsymbol{C}_0=(c_{ij},i,j=1,\cdots,q)$, $\boldsymbol{c}=\otimes^s\boldsymbol{c}_0$, $\boldsymbol{C}=\otimes^s\boldsymbol{C}_0$($\otimes$ 是 Kronecker 乘积),

$$c_i=1+\left|\frac{2i-1-q}{4q}\right|-\frac{(2i-1-q)^2}{8q^2}$$

和

$$c_{ij}=1+\left|\frac{2i-1-q}{4q}\right|+\left|\frac{2j-1-q}{4q}\right|-\left|\frac{i-j}{2q}\right|.$$

公式 (6.2) 在理论研究中的作用很大. 如果因子所有的水平组合出现相同的次数, 那么把这样的设计称为*全面的* (或者*完全的*). 任何的完全设计是一个正交表, 整数 $\boldsymbol{y}(\mathcal{P})$ 所对应的向量是 $\mathbf{1}$ 的倍数. 对任何的因子设计 $\mathcal{P}\in\mathcal{D}(n;q^s)$, $y(\mathcal{P})/n$ 可以认为是在 q^s 水平组合下的一个测量. 因此, 可以将 $\boldsymbol{y}$ 扩展为一个 q^s 维的正值向量, 并且满足 $\boldsymbol{y}^{\mathrm{T}}\mathbf{1}=n$. 在这些符号的定义下, Ma 等 (2003) 得到了以下结果.

定理 6.3　*令 $\mathcal{P}\in\mathcal{D}(n;q^s)$ 是一个包含 n 个格子点的集合. 那么,*

(1) 当 $q=2$ 或是奇数时, $\mathcal{P}$ 的 $\mathrm{CD}(\mathcal{P})$ 值在 $\mathcal{D}(n;q^s)$ 上达到极小当且仅当 $\boldsymbol{y}(\mathcal{P})=(n/q^s)\mathbf{1}$;

(2) 当 q 是偶数 (但不等于 2) 时, $\mathcal{P}$ 的 $\mathrm{CD}(\mathcal{P})$ 值在 $\mathcal{D}(n;q^s)$ 上达到极小当且仅当

$$\boldsymbol{y}(\mathcal{P})=\frac{n}{q^s}\otimes^s\begin{pmatrix}\mathbf{1}_{q/2-1}\\1-\dfrac{1}{4(4q+1)}\\1-\dfrac{1}{4(4q+1)}\\\mathbf{1}_{q/2-1}\end{pmatrix}+n\frac{1-\left(1-\dfrac{1}{2q(4q+1)}\right)^s}{2^s}\otimes^s\begin{pmatrix}\mathbf{0}_{q/2-1}\\1\\1\\\mathbf{0}_{q/2-1}\end{pmatrix},$$

其中 $\otimes$ 是 Kronecker 乘积;

(3) 当 q 是偶数 (但不等于 2) 且 $\mathcal{P}\in\mathcal{U}(n;q^s)$ 时, $\mathcal{P}$ 的 $\mathrm{CD}(\mathcal{P})$ 在 $\mathcal{D}(n;q^s)$ 上达到极小当且仅当

$$\boldsymbol{y}(\mathcal{P})=n\boldsymbol{C}^{-1}\boldsymbol{c}-n\left[\frac{\mathbf{1}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{c}}{\mathbf{1}^{\mathrm{T}}\boldsymbol{C}^{-1}\mathbf{1}}\left(1-\frac{s(8q^2+2q)}{8q^2+2q-1}\right)+\frac{s-1}{\mathbf{1}^{\mathrm{T}}\boldsymbol{C}^{-1}\mathbf{1}}\right]\boldsymbol{C}^{-1}\mathbf{1}$$

$$
-\frac{n}{q}\frac{(\mathbf{1}^{\mathrm{T}}\boldsymbol{C}_0^{-1}\boldsymbol{c}_0)^{s-1}-1}{(\mathbf{1}^{\mathrm{T}}\boldsymbol{C}_0^{-1}\mathbf{1})^{s-1}}\sum_{i=1}^{s}(\boldsymbol{C}_0^{-1}\mathbf{1})^{i-1}\otimes\mathbf{1}\otimes(\boldsymbol{C}_0^{-1}\mathbf{1})^{i-1}.
$$

特别地, 当 $s=1,2$ 时, 有 $\boldsymbol{y}(\mathcal{P})=(n/q^s)\mathbf{1}$.

该定理的具体证明可以参见 (Ma et al., 2003).

定理 6.3 说明完全设计是 CD 意义下的均匀设计, 并且当 $\mathcal{P}$ 是完全设计时猜想是正确的. 这也说明了当 $q=2$ 且 $n=2^{s-1}$ 时, 猜想也是正确的, 此时后者不是一个完全设计. 在 q 是偶数的情形下, 猜想是很复杂的. 例如, 当 $q=4$ 且 $s=2$ 时, 均匀设计的 $n(i,j)$ 可以写为

$$
\boldsymbol{N}\equiv(n(i,j))=\frac{n}{16}\begin{pmatrix} 1 & \frac{67}{68} & \frac{67}{68} & 1 \\ \frac{67}{68} & \frac{70}{68} & \frac{70}{68} & \frac{67}{68} \\ \frac{67}{68} & \frac{70}{68} & \frac{70}{68} & \frac{67}{68} \\ 1 & \frac{67}{68} & \frac{67}{68} & 1 \end{pmatrix}.
$$

$\boldsymbol{N}$ 的所有元素接近于 $\frac{n}{16}$. 因为 (6.2) 中目标函数是 $\boldsymbol{y}$ 的连续函数, 所以最优的 $\boldsymbol{y}$-值是 $\mathcal{D}(n;q^s)$ 中与 $\boldsymbol{N}$ 有着最小距离的设计. 当 n 是 16 的除数且不太大时, 得到的结果是 $L_n(4^5)$. 该事实说明当 q 是偶数且把 CD 当作均匀测度时, 猜想不一定正确.

当格子设计 $\mathcal{P}$ 不是完全的, 猜想是不是依旧成立? 注意到 $[\mathrm{CD}(\mathcal{P})]^2=\sum_{u\neq\varnothing}I_u(\mathcal{P})^2$, 其中 $I_u(\mathcal{P})=I_{|u|}(\mathcal{P}^u)$, $\mathcal{P}^u$ 是 $\mathcal{P}$ 在 C^u 上的投影, 且

$$
I_u(\mathcal{P})^2=\int_{C^u}\left(\frac{N(\mathcal{P}^u,J_{\boldsymbol{x}^u})}{n}-\mathrm{Vol}(J_{\boldsymbol{x}^u})\right)^2\mathrm{d}\boldsymbol{x}^u.
$$

定义一个新的均匀性测度

$$
[\mathrm{CD}_{2,t}(\mathcal{P})]^2=\sum_{0<|u|\leqslant t}I_u(\mathcal{P})^2.
$$

Ma 等 (2003) 证明了如下定理.

定理 6.4 在 $\mathrm{CD}_{2,t}$ 测度下的均匀设计 $U_n(q^s)$ 是正交表 $\mathrm{OA}(n;q^s;t)$ (如果存在这样的正交表), 其中 $t<s, n$ 是 q^t 的倍数, q 等于 2 或奇数.

Liu (2002) 用离散偏差 (2.5.1 小节) 均匀性测度对该猜想进行了研究. Liu (2002) 和 Fang 等 (2003c) 证明了对称饱和正交表是离散偏差意义下的均匀设计. 同时, 按照 (Li et al., 2004) 和 (Liu et al., 2006) 的结果, 我们知道根据 $j=2$ 时

(6.25) 定义的二维投影偏差, 非对称饱和正交表大多数也是均匀的. Tang (2005) 在他的博士学位论文中对该猜想做了进一步的讨论.

Fang 等 (2002c) 从其他的角度对上面的研究提出了另一种非正交性的测度. 对于一个设计 $\mathcal{P} \in \mathcal{D}(n; q^s)$, $\boldsymbol{y}(\mathcal{P})$ 和以前的定义一样是个 q^s 维向量. 在下面的讨论中令 $v = q^s$, $\boldsymbol{e}$ 和 $\boldsymbol{I}$ 分别表示所有元素相同的 $q \times 1$ 向量和 $q \times q$ 单位矩阵, 显然 $\boldsymbol{J} = \boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}$. 用 Ω 来表示 0 和 1 的二进制 s 元组的集合. 记

$$\boldsymbol{L}(0) = q^{-1}\boldsymbol{J}, \quad \boldsymbol{L}(1) = \boldsymbol{I} - q^{-1}\boldsymbol{J}, \quad \boldsymbol{G}(0) = \boldsymbol{e}^{\mathrm{T}}, \quad \boldsymbol{G}(1) = \boldsymbol{I}. \tag{6.3}$$

$\boldsymbol{e}$, $\boldsymbol{I}$ 和 $\boldsymbol{J}$ 的 t 重 Kronecker 乘积分别表示为 $\boldsymbol{e}^{(t)}$, $\boldsymbol{I}^{(t)}$ 和 $\boldsymbol{J}^{(t)}$. 对任何的 $\boldsymbol{x} = x_1 \cdots x_s \in \Omega$, 定义矩阵

$$\boldsymbol{W}(\boldsymbol{x}) = \boldsymbol{L}(x_1) \otimes \cdots \otimes \boldsymbol{L}(x_s), \quad \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{G}(x_1) \otimes \cdots \otimes \boldsymbol{G}(x_s), \tag{6.4}$$

它们的阶数分别为 $v \times v$ 和 $q^{\sum x_i} \times v$. 这里 $\otimes$ 是 Kronecker 乘积. 对 $0 \leqslant i \leqslant s$, Ω_i 是恰好由 i 个单位元组成的二进制 s- 元组 Ω 的子集. 同时, $\Omega^* = \Omega - \Omega_0$ 表示 Ω 中非空元素的集合. 由 (6.3) 和 (6.4), 对 $\boldsymbol{x} \in \Omega^*$, 矩阵 $W(\boldsymbol{x})$ 的累积等于 $\boldsymbol{I}^{(s)} - v^{-1}\boldsymbol{J}^{(s)}$. 因此

$$\boldsymbol{y}(\mathcal{P})^{\mathrm{T}}\{\boldsymbol{I}^{(s)} - v^{-1}\boldsymbol{J}^{(s)}\}\boldsymbol{y}(\mathcal{P}) = \sum_{\boldsymbol{x} \in \Omega^*} \boldsymbol{y}(\mathcal{P})^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{x})\boldsymbol{y}(\mathcal{P}). \tag{6.5}$$

很容易看到矩阵 $\boldsymbol{W}(\boldsymbol{x})$ 是幂等的, 且对每一个 $\boldsymbol{x} \neq \boldsymbol{y}$, $\boldsymbol{W}(\boldsymbol{x})\boldsymbol{W}(\boldsymbol{y}) = \boldsymbol{O}$. 这样等式 (6.5) 的右端是左端的一个正交分解, 同时, 等式左端可以进一步解释如下. 将 s 个因子定义为 $F_1, \cdots, F_s$. 对任何 $\boldsymbol{x} = x_1 \cdots x_s \in \Omega^*$, 令 $F(\boldsymbol{x})$ 表示 $F^{x_1} \cdots F^{x_s}$ 的交互效应; 通常, 一个因子交互效应就是该因子的主效应. 现在考虑每一个 $v = q^s$ 水平组合恰好出现一次的假设全因子. 如果一个来源于假设全因子并按字典序排列的观察值的 $v \times 1$ 向量等于 $\boldsymbol{y}(\mathcal{P})$, 那么通过 (6.3) 和 (6.4), 由交互效应 $F(\boldsymbol{x})$ 产生的平方和可以表示为

$$\mathrm{SS}(\boldsymbol{x}) = \boldsymbol{y}(\mathcal{P})^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{x})\boldsymbol{y}(\mathcal{P}). \tag{6.6}$$

这样 (6.5) 式就是一个方差分解的表达式. 根据 (6.6), 可以推断出对每个 $\boldsymbol{x} \in \Omega_1 \bigcup \cdots \bigcup \Omega_t$, $\mathrm{SS}(\boldsymbol{x})$ $(= \boldsymbol{y}(\mathcal{P})^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{x})\boldsymbol{y}(\mathcal{P}))$ 不存在当且仅当设计 $\mathcal{P}$ 是一个强度为 $t(1 \leqslant t \leqslant s)$ 的正交表. 因此, 记

$$B_i(\mathcal{P}) = \sum_{\boldsymbol{x} \in \Omega_i} \boldsymbol{y}(\mathcal{P})^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{x})\boldsymbol{y}(\mathcal{P}), \quad 1 \leqslant i \leqslant s, \tag{6.7}$$

用强度为 t 的正交表来表示的 $\mathcal{P}$ 的偏离可以用 $\sum_{i=1}^{t} B_i(\mathcal{P})$ 来测量, 这样 $B_1(d)$ 是度量 $\mathcal{P}$ 偏离一个强度为 1 的正交表, $B_2(d)$ 是度量 $\mathcal{P}$ 偏离一个强度为 2 的正交表,

以此类推. 换句话说, $B_1(\mathcal{P}),\cdots,B_s(\mathcal{P})$ 考察了 $\mathcal{P}$ 偏离强度逐渐更高的正交表的程度. 因此, 为了保证接近强度相继增高的正交表, 可以选择使得 $B_1(\mathcal{P})$, $B_2(\mathcal{P})$, $\cdots$ 依次最小的 $\mathcal{P}$. 这类似于有人在正规部分因子设计中使用的最小低阶混杂准则. 对一个行数为 $n=q^{s-p}$ 的正规部分因子设计 $\mathcal{P}$, Fang 等 (2002c) 发现

$$B_i(\mathcal{P})=\frac{n^2(q-1)}{q^s}A_i(\mathcal{P}),\quad 1\leqslant i\leqslant s, \tag{6.8}$$

其中 $(A_1(\mathcal{P}),\cdots,A_s(\mathcal{P}))$ 是 $\mathcal{P}$ 的字长型. 该结果说明了正规部分混杂与正交性之间的联系. 在 6.4 节中, 我们介绍一些混杂与均匀的联系. 紧接着通过 (6.8) 式, 可以建立正交性与均匀性的一些联系, 具体可以参见 (Fang et al., 2002c).

6.3 均匀性与混杂

两个同构的因子设计在 ANOVA 模型中的统计推断能力是一样的, 在这意义下我们认为它们是等价的. 然而, 两个同构的设计却可能有着不同的均匀性. 例如, 表 6.4 中的两个 $L_9(3^4)$ 设计 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$, 它们之间是同构的, 但是它们的 CD 值分别是 0.050059 和 0.0493645. 其中, 设计 $\boldsymbol{U}_1$ 可以很容易地在文献中找到, 而设计 $\boldsymbol{U}_2$ 是在 $\mathcal{U}(9;3^4)$ 类上通过最小化 CD 值得到的 (参见 (Fang and Winker, 1998)).

表 6.4 两个 $L_9(3^4)$ 设计

编号	$\boldsymbol{U}_1$				$\boldsymbol{U}_2$			
1	1	1	1	1	1	1	1	2
2	1	2	2	2	1	2	3	1
3	1	3	3	3	1	3	2	3
4	2	1	2	3	2	1	3	3
5	2	2	3	1	2	2	2	2
6	2	3	1	2	2	3	1	1
7	3	1	3	2	3	1	2	1
8	3	2	1	3	3	2	1	3
9	3	3	2	1	3	3	3	2

假设在一个试验中有三个三水平因子 A,B 和 C. 我们可以从设计 $\boldsymbol{U}_1$ 或者 $\boldsymbol{U}_2$ 中任选三列来安排因子. 现分别将由 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$ 的前三列生成的设计记为 $\mathcal{P}_1$ 和 $\mathcal{P}_2$. 把因子 A 的主效应分解为线性和二次两个成分, 分别记为 A_l 和 A_q(类似地可以定义 B_l,B_q, C_l 和 C_q). 如果因子 A 与 B 之间的交互效应 $A\times B$ 存在, 可以将其分解为四个成分 A_lB_l,A_lB_q, A_qB_l 和 A_qB_q (Fang and Ma, 2000; Box and Draper, 1987, 236-239 页) 当在试验中存在着交互效应 $A\times B$, $A\times C$ 和 $B\times C$ 时, 若想正确地将其与主效应分开是不可能的. 利用 $\mathcal{P}_2$, 可得它们的主效应和它们之

间的交互效应混杂情形如下:

$$\begin{aligned} A_l &= 0.5B_lC_q + 0.5B_qC_l, \\ A_q &= 1.5B_lC_l - 0.5B_qC_q, \\ B_l &= 0.5A_lC_q + 0.5A_qC_l, \\ B_q &= 1.5A_lC_l - 0.5A_qC_q, \\ C_l &= 0.5A_lB_q + 0.5A_qB_l, \\ C_q &= 1.5A_lB_l - 0.5A_qB_q. \end{aligned}$$

若用 $\mathcal{P}_1$, 其混杂情形为

$$\begin{aligned} A_l &= -0.75B_lC_l - 0.25B_lC_q + 0.25B_qC_l - 0.25B_qC_q, \\ A_q &= 0.75B_lC_l - 0.75B_lC_q + 0.75B_qC_l + 0.25B_qC_q, \\ B_l &= -0.75A_lC_l - 0.25A_lC_q + 0.25A_qC_l - 0.25A_qC_q, \\ B_q &= 0.75A_lC_l - 0.75A_lC_q + 0.75A_qC_l + 0.25A_qC_q, \\ C_l &= -0.75A_lB_l + 0.25A_lB_q + 0.25A_qB_l + 0.25A_qB_q, \\ C_q &= -0.75A_lB_l - 0.75A_lB_q - 0.75A_qB_l + 0.25A_qB_q. \end{aligned}$$

如果将高阶交互效应 $A_lB_q, A_qB_l, A_qB_q, \cdots, B_qC_q$ 忽略, $\mathcal{P}_2$ 的混杂情形变为

$$\mathcal{P}_2: \quad \begin{cases} A_q = 1.5B_lC_l, \\ B_q = 1.5A_lC_l, \\ C_q = 1.5A_lB_l. \end{cases}$$

在这种情况下, 可以估计出所有的线性效应 A_l, B_l 和 C_l, 它们是完全没有混杂的. 而对 $\mathcal{P}_1$, 有

$$\mathcal{P}_1: \quad \begin{cases} A_l = -0.75B_lC_l, \quad A_q = 0.75B_lC_l, \\ B_l = -0.75A_lC_l, \quad B_q = 0.75A_lC_l, \\ C_l = -0.75A_lB_l, \quad C_q = -0.75A_lB_l. \end{cases}$$

在这种情况下, 主效应都与线性交互效应混杂. 显然, 在混杂的意义下, 设计 $\mathcal{P}_2$ 是优于 $\mathcal{P}_1$ 的. 从 $\boldsymbol{U}_1$ 或者 $\boldsymbol{U}_2$ 中任取三列是有四种选择的. 其中, 在 $\boldsymbol{U}_1$ 的四种选择中只有一种是与 $\mathcal{P}_2$ 的混杂情形是一样的, 而其他三种选择是与 $\mathcal{P}_1$ 有相同的混杂情形. 另一方面, 对于 $\boldsymbol{U}_2$ 的所有四种选择均与 $\mathcal{P}_2$ 的混杂情形一样. 因此, 得出的结论是, 在混杂的意义下, $\boldsymbol{U}_2$ 是优于 $\boldsymbol{U}_1$ 的. 通过该例子, Fang 和 Ma (2000) 提出了下面的概念.

定义 6.3　对于给定的 (n, q, s), 一个正交设计 $L_n(q^s)$ 称为均匀正交设计, 如果它的 CD 值是所有这类正交设计中最小的, 记为 $\mathrm{UL}_n(q^s)$.

显然, 我们可以用均匀性的其他测度来代替定义 6.3 中的 CD. Fang 和 Winker (1998) 已经得到 $q > 2$ 的几种均匀正交设计. 这些设计的性质仍需要研究. Hickernell 和 Liu (2002) 利用了再生核的方法说明了均匀设计限制了别名的效应来产生合理的效率和稳健性. 最近, Tang 等 (2012) 提出了水平置换的方法 (见 3.5 节) 和均匀最小低阶混杂 (UMA) 设计的概念, 并且得到了 CD 准则下的一个 UMA 设计 $U_{27}(3^{13})$. 随后, Fang 等 (2016) 找到了 MD 测度下的一个 UMA 设计 $U_{27}(3^{13})$, 并得到 $\mathrm{MD}(\mathrm{UL}_{27}) = 62.8011$, 该结果可见表 6.5. 另外此设计也是一个强度为 2 的正交表, 并且前 m 列可构成最小低阶混杂设计, 其中 $m = 4, 5, 6, 10, 11, 12, 13$.

表 6.5　MD 测度下的一个均匀最小低阶混杂设计 $U_{27}(3^{13})$

编号	1	2	3	4	5	6	7	8	9	10	11	12	13
1	2	2	2	2	3	2	1	2	3	1	3	3	1
2	2	2	3	2	3	3	2	3	1	3	2	2	3
3	2	2	1	2	3	1	3	1	2	2	1	1	2
4	2	3	2	3	1	2	3	3	2	1	2	2	2
5	2	3	3	3	1	3	2	2	3	2	1	1	1
6	2	3	1	3	1	1	1	1	1	3	3	3	3
7	2	1	2	1	2	2	3	2	1	3	1	1	3
8	2	1	3	1	2	3	1	3	2	2	3	3	2
9	2	1	1	1	2	1	2	1	3	1	2	2	1
10	3	2	2	1	1	3	3	1	1	2	3	2	1
11	3	2	3	1	1	1	1	2	2	1	2	1	3
12	3	2	1	1	1	2	2	3	3	3	1	3	2
13	3	3	2	2	2	3	2	1	2	1	1	3	3
14	3	3	3	2	2	1	3	2	3	3	3	2	2
15	3	3	1	2	2	2	1	3	1	2	2	1	1
16	3	1	2	3	3	3	1	1	3	3	2	1	2
17	3	1	3	3	3	1	3	3	1	1	1	3	1
18	3	1	1	3	3	2	2	2	2	2	3	2	3
19	1	2	2	3	2	1	1	3	3	2	1	2	3
20	1	2	3	3	2	2	2	1	1	1	3	1	2
21	1	2	1	3	2	3	3	2	2	3	2	3	1
22	1	3	2	1	3	1	2	2	1	2	2	3	2
23	1	3	3	1	3	2	1	1	2	3	1	2	1
24	1	3	1	1	3	3	3	3	3	1	3	1	3
25	1	1	2	2	1	1	2	3	2	3	3	1	1
26	1	1	3	2	1	2	3	1	3	2	2	3	3
27	1	1	1	2	1	3	1	2	1	1	1	2	2

6.4 均匀性与字长型

基于字长型的两个重要准则是分辨度和最小低阶混杂准则 (相关定义见 1.4 节). 这些准则表面看起来与均匀性准则是毫不相关的, 直到 Fang 和 Mukerjee (2000) 发现了二水平的因子设计的均匀性与字长型之间有解析关系. 他们得到的主要结果如下.

定理 6.5 若设计 $\boldsymbol{U} \in \mathcal{D}(2^{s-p}; 2^s)$ 是正规的, 即 $\boldsymbol{U}$ 是正规部分因子设计 2^{s-p}, 则

$$\mathrm{CD}^2(\boldsymbol{U}) = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \left(\frac{9}{8}\right)^s \left\{1 + \sum_{i=1}^{s} \frac{A_i(\boldsymbol{U})}{9^i}\right\}, \tag{6.9}$$

其中 $(A_1(\boldsymbol{U}), \cdots, A_s(\boldsymbol{U}))$ 为设计 $\boldsymbol{U}$ 的字长型.

从等式中可以看出对于正规因子设计 2^{s-p}, 最小低阶混杂准则和 CD 均匀性测度基本上是等价的. 若通过 CD 值来比较两个设计 $\boldsymbol{U}_1$ 和 $\boldsymbol{U}_2$, 等价于比较 $\sum_{i=1}^s \frac{A_i(\boldsymbol{U}_1)}{9^i}$ 和 $\sum_{i=1}^s \frac{A_i(\boldsymbol{U}_2)}{9^i}$. 若设计 $\boldsymbol{U}_1$ 的分辨度 (记为 t) 比设计 $\boldsymbol{U}_2$ 的分辨度 (记为 t') 大, 有

$$\sum_{i=1}^{s} \frac{A_i(\boldsymbol{U}_1)}{9^i} = \sum_{i=t}^{s} \frac{A_i(\boldsymbol{U}_1)}{9^i}$$

和

$$\sum_{i=1}^{s} \frac{A_i(\boldsymbol{U}_2)}{9^i} = \sum_{i=t'}^{s} \frac{A_i(\boldsymbol{U}_2)}{9^i} = \sum_{i=t'}^{t-1} \frac{A_i(\boldsymbol{U}_2)}{9^i} + \sum_{i=t}^{s} \frac{A_i(\boldsymbol{U}_2)}{9^i}.$$

显然, 由于 $A_i(\cdot)/9^i$ 中 $A_i(\cdot)$ 的系数随着 i 的增大而呈指数减少, 所以 $\boldsymbol{U}_1$ 更有可能比 $\boldsymbol{U}_2$ 有更小的 CD 值. 马长兴检验了 (Chen et al., 1993) 中所有的二水平正规设计, 发现通过最小低阶混杂准则和 CD 均匀性测度所选出来的设计是一致的.

等式 (6.9) 适用于正规二水平因子设计. 在过去的十几年里, 对于该结果有一些推广. Ma 和 Fang (2001) 将这种联系推广到在 CD, WD 和对称 L_2 偏差下的更一般的二水平因子设计中, 以及在 WD 下的三水平设计中. Fang 和 Ma (2002) 将这种联系推广至正规部分因子设计 3^{s-1} 中, 并发现在 CD 下若想得到更一般的结果是有难度的. Ye (2003) 和 Sun 等 (2011a) 利用示性函数证实了不管是正规的还是非正规的或者有无重复的所有的二水平因子设计均是满足 (6.9) 式的. Sun 等 (2011a) 将定理 6.5 中的结果推广到在 WD 下的更一般二水平和三水平因子设计中. 下面的定理给出了他们的结果.

定理 6.6 对于任意的 $\mathcal{P} \in \mathcal{D}(n; q^s)(q = 2$ 或 $3)$,

$$
\mathrm{WD}^2(\mathcal{P}) = \begin{cases} -\left(\dfrac{4}{3}\right)^s + \left(\dfrac{11}{8}\right)^s \left[1 + \displaystyle\sum_{r=1}^{s} \dfrac{A_r(\mathcal{P})}{11^r}\right], & q = 2, \\ -\left(\dfrac{4}{3}\right)^s + \left(\dfrac{73}{54}\right)^s \left[1 + \displaystyle\sum_{r=1}^{s} \left(\dfrac{4}{73}\right)^r A_r(\mathcal{P})\right], & q = 3, \end{cases} \tag{6.10}
$$

其中 $(A_1(\mathcal{P}), \cdots, A_s(\mathcal{P}))$ 是 (1.40) 中定义的广义字长型.

定理 6.6 的证明是基于 (4.8) 式中 $\mathrm{WD}^2(\mathcal{P})(q_1 = \cdots = q_s)$ 的二次型, Cheng 和 Ye (2004) 重新定义的广义字长型, 以及一些矩阵计算. 具体的细节这里就不赘述了, 有兴趣的读者可参见 (Sun et al., 2011a).

注意到基于广义字长型 $(A_1^g(\mathcal{P}), \cdots, A_s^g(\mathcal{P}))$ 和 CD 均匀性测度, 这些联系与 (Ma and Fang, 2001) 中所提的是一致的, 但是 (Ma and Fang, 2001) 中只证明了 $q = 2$ 时正规的情形, 对一般的情形并没有提及. (Ma and Fang, 2001) 中所提到的几个例子也说明了因子设计中的这些联系能明显降低计算复杂度, 另外也提供了利用均匀性来搜索最小低阶混杂设计的一种方法.

从定理 6.5 和定理 6.6, 可以注意到 $\mathrm{WD}^2(\mathcal{P})$ 或 $\mathrm{CD}^2(\mathcal{P})$ 中 $A_r(\mathcal{P})$ 的系数是随着 r 的增大而呈指数减少的, 所以有更小低阶混杂的设计往往会有着更小的 $\mathrm{WD}^2(\mathcal{P})$ 或 $\mathrm{CD}^2(\mathcal{P})$. 在 $\mathrm{WD}^2(\mathcal{P})$ 或 $\mathrm{CD}^2(\mathcal{P})$ 测度下的均匀设计与 GMA 设计有着密切的关系. 事实上, Sun 等 (2011a) 得到的下面的定理就说明了在哪些条件下可以保证 $[\mathrm{WD}(\mathcal{P})]^2$ 或 $[\mathrm{CD}(\mathcal{P})]^2$ 与 GMA 之间的一致性.

定理 6.7 (i) 假设 $\mathcal{P}_1, \mathcal{P}_2 \in \mathcal{D}(n; q^s)$, k 为常数, 当 $r = 1, \cdots, s$, $i = 1, 2$ 时, $[kA_r(\mathcal{P}_i)]$为整数, 则

(1) 若 $q = 2$ 且 $\max\{kA_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 8$, 则 $\mathrm{CD}^2(\mathcal{P}_1) < \mathrm{CD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂;

(2) 若 $q = 2$ 且 $\max\{kA_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 10$, 则 $\mathrm{WD}^2(\mathcal{P}_1) < \mathrm{WD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂;

(3) 若 $q = 3$ 且 $\max\{kA_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 69/4$, 则 $\mathrm{WD}^2(\mathcal{P}_1) < \mathrm{WD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂.

(ii) 假设 $\mathcal{P}_1, \mathcal{P}_2 \in \mathcal{D}(n; q^s)$ 且当 $r < t$ 时, 存在着正整数 t 使得 $A_t(\mathcal{P}_1) \leqslant A_t(\mathcal{P}_2) - 1$, $A_r(\mathcal{P}_1) = A_r(\mathcal{P}_2)$, 则

(1) 若 $q = 2$ 且 $\max\{A_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 8$, 则 $\mathrm{CD}^2(\mathcal{P}_1) < \mathrm{CD}^2(\mathcal{P}_2)$;

(2) 若 $q = 2$ 且 $\max\{A_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 10$, 则 $\mathrm{WD}^2(\mathcal{P}_1) < \mathrm{WD}^2(\mathcal{P}_2)$;

(3) 若 $q = 3$ 且 $\max\{A_r(\mathcal{P}_i), r = 1, \cdots, s, i = 1, 2\} \leqslant 69/4$, 则 $\mathrm{WD}^2(\mathcal{P}_1) < \mathrm{WD}^2(\mathcal{P}_2)$.

定理 6.7 的证明与定理 6.10 是类似的, 即根据 (6.9) 式, (6.10) 式和下面的引

理可直接得到.

引理 6.8　假设 a_i 和 b_i 均是非负数, 并且当 $i=0,\cdots,k$ 时, $a_i, b_i \leqslant m-1$.

(i) 若 $a_k \leqslant b_k - 1$, 则 $\sum_{i=0}^k a_i m^i < \sum_{i=0}^k b_i m^i$.

(ii) 若 a_i 和 b_i 都是整数并且 $a_k \neq b_k$, 则

$$\sum_{i=0}^k a_i m^i < \sum_{i=0}^k b_i m^i \text{ 的充要条件是 } a_k < b_k.$$

证明　(i) 可由下式得到

$$\begin{aligned}\sum_{i=0}^k a_i m^i &= a_k m^k + \sum_{i=0}^{k-1} a_i m^i \leqslant a_k m^k + (m-1)\sum_{i=0}^{k-1} m^i \\ &= a_k m^k + m^k - 1 < b_k m^k \leqslant \sum_{i=0}^k b_i m^i.\end{aligned}$$

(ii) 可由 (i) 直接得到. 证毕.

如果用 (2.53) 式定义的离散偏差作为均匀性测度, Qin 和 Fang (2004) 得到了均匀性与 GMA 之间的关系如下.

定理 6.9　对于任意的 $\mathcal{P} \in \mathcal{D}(n; q^s)$,

$$\mathrm{DD}^2(\mathcal{P}) = \left(\frac{a+(q-1)b}{q}\right)^s \sum_{j=1}^s \left(\frac{a-b}{a+(q-1)b}\right)^j A_j(\mathcal{P}),$$

其中 $\mathrm{DD}(\mathcal{P})$ 和 $A_j(\mathcal{P})$ 分别在 (2.53) 与 (1.40) 中有所定义.

从上面的联系中可以看出, $\mathrm{DD}(\mathcal{P})$ 中 $A_j(\mathcal{P})$ 的系数是随着 j 的增大而呈指数减少的, 由此我们推断那些 j 较小时 $A_j(\mathcal{P})$ 也较小, 即那些有着更小低阶混杂的设计在 DD 意义下的表现也较好. 这说明了在离散偏差下的均匀设计与 GMA 设计有着密切的关系, 并为 GMA 准则提供了在离散偏差下的解释. 定理 6.9 说明了 (2.53) 式的最小离散偏差与 GMA 准则并不是完全的一致. 那这两种准则在什么条件下等价呢? Sun 等 (2011a) 更直觉地得到了定理 6.9, 并给出了二者等价的一个条件如下.

定理 6.10　假设 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 是 $\mathcal{D}(n; q^s)$ 中没有重复的两个设计. 如果

$$\frac{a+(q-1)b}{a-b} - 1 = \frac{qb}{a-b} \geqslant n(q^s - n),$$

则 $\mathrm{DD}^2(\mathcal{P}_1) < \mathrm{DD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂.

证明　由 (Hickernell and Liu, 2002) 的定理 2 有

$$A_r(\mathcal{P}) = \frac{1}{n^2} \sum_{|\boldsymbol{u}|=r} \sum_{i,k=1}^n \prod_{l\in\boldsymbol{u}} (-1 + q\delta_{a_{il}a_{kl}}), \tag{6.11}$$

其中 $\boldsymbol{u}$ 是 $\{1,\cdots,s\}$ 中的子集, $|\boldsymbol{u}|$ 表示集合 $\boldsymbol{u}$ 中的元素个数, $\mathcal{P}=(a_{ij})$, 因此 $n^2A_r(\mathcal{P})$ 是一个整数. 此外, 由 (Cheng and Ye, 2004) 中的定理 4.1, 可知对于任何没有重复的 $\mathcal{P}\in\mathcal{D}(n;q^s)$, 有

$$\sum_{r=1}^{s}A_r(\mathcal{P})=\frac{q^s}{n}-1.$$

所以, 对于任意的 r, 有 $n^2A_r(\mathcal{P})\leqslant n(q^s-n)$. 于是由定理 6.9 和引理 6.8, 结论易得.

由该定理, 可以得知当存在没有重复的正交设计 $L(n;q^s)$ 并且 $qb/(a-b)\geqslant n(q^s-n)$ 成立时, 在 DD 意义下的均匀设计是正交的, 并且也是所有没有重复的 $\mathcal{D}(n;q^s)$ 中的 GMA 设计. 更进一步, Sun 等 (2011a) 通过引理 6.8 得到了在哪些条件下最小离散偏差与 GMA 是一致的:

定理 6.11 (i) 假设 $\mathcal{P}_1,\mathcal{P}_2\in\mathcal{D}(n;q^s)$, k 为常数, 当 $r=1,\cdots,s$, $i=1,2$ 时, 所有的 $[kA_r(\mathcal{P}_i)]$ 是整数. 若 $qb/(a-b)\geqslant\max\{kA_r(\mathcal{P}_i),r=1,\cdots,s,i=1,2\}$, 则 $\mathrm{DD}^2(\mathcal{P}_1)<\mathrm{DD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂.

(ii) 假设 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 是 $\mathcal{D}(n;q^s)$ 中的两个正规设计. 若 $qb/(a-b)\geqslant\max\{A_r(\mathcal{P}_i),r=1,\cdots,s,i=1,2\}$, 则 $\mathrm{DD}^2(\mathcal{P}_1)<\mathrm{DD}^2(\mathcal{P}_2)$ 等价于 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的低阶混杂.

(iii) 假设 $\mathcal{P}_1,\mathcal{P}_2\in\mathcal{D}(n;q^s)$. 若 $qb/(a-b)\geqslant\max\{A_r(\mathcal{P}_i),r=1,\cdots,s,i=1,2\}$ 并且存在正整数 t 使得, 当 $r<t$ 时, $A_r(\mathcal{P}_1)=A_r(\mathcal{P}_2)$, $A_t(\mathcal{P}_1)\leqslant A_t(\mathcal{P}_2)-1$, 且 $\mathrm{DD}^2(\mathcal{P}_1)<\mathrm{DD}^2(\mathcal{P}_2)$.

此外, Hickernell 和 Liu (2002) 也说明了 GMA 设计在某些条件下与均匀设计是等价的, 相关结果可见 6.5.1 小节的定理 6.15. Hickernell 和 Liu (2002) 与 Fang 和 Qin (2004) 研究了均匀设计的投影性与低阶混杂之间的关系, 具体可见 6.5 节.

下面考虑任意水平的任意测度下的均匀性准则与 GMA 准则之间的关系. 我们用 (n,q^s) 来表示 n 行, s 个因子且所有因子的水平均是 q 的设计. 对于一个 (n,q^s) 设计 $\mathcal{P}$, 每个因子都有 $q!$ 个可能的水平置换, 那么总共就有 $(q!)^s$ 个可能的组合同构设计. 用 $\mathcal{H}(\mathcal{P})$ 来表示这些组合同构设计构成的集合. 由于行置换与列置换不影响设计的几何结构和统计性质, 所以这里不需要考虑设计的行列置换. 对于一个有 n 个点的设计 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$ 和一个非负函数 $F(\boldsymbol{x}_i,\boldsymbol{x}_j)\geqslant 0$, 定义

$$\phi(\mathcal{P})=\frac{1}{n^2}\sum_{i,j=1}^{n}F(\boldsymbol{x}_i,\boldsymbol{x}_j). \tag{6.12}$$

称 $\phi(\mathcal{P})$ 为基于 F 的 $\mathcal{P}$ 的一个空间填充测度. $\mathcal{H}(\mathcal{P})$ 中所有设计的 GMP 是相同的, 但是 $\phi(\mathcal{P})$ 并不是完全一样的. 我们计算 $\mathcal{H}(\mathcal{P})$ 中每个设计的 $\phi(\mathcal{P})$ 和所有设计的总均值 $\bar{\phi}(\mathcal{P})$, 并定义

$$\bar{\phi}(\mathcal{P})=\frac{1}{(q!)^s}\sum_{\mathcal{P}'\in\mathcal{H}(\mathcal{P})}\phi(\mathcal{P}'). \tag{6.13}$$

下面的定理说明了 (6.13) 式中的 $\bar{\phi}(\mathcal{P})$ 可以在一类广泛的空间填充测度下表示成 GMP 的线性组合.

引理 6.12　假设 $F(\boldsymbol{x}_i, \boldsymbol{x}_j) = \prod_{k=1}^{s} f(x_{ik}, x_{jk})$, 且 $f(\cdot,\cdot)$ 满足

$$\begin{cases} f(x,x)+f(y,y)>f(x,y)+f(y,x), \\ f(x,y) \geqslant 0, \quad \forall\, x \neq y,\ x,y \in [0,1]. \end{cases} \tag{6.14}$$

对于一个 $(n;q^s)$ 设计 $\mathcal{P}$, 有

$$\bar{\phi}(\mathcal{P}) = \left(\frac{c_1(c_2+q-1)}{q^2(q-1)}\right)^s \sum_{i=0}^{s} \left(\frac{c_2-1}{c_2+q-1}\right)^i A_i(\mathcal{P}), \tag{6.15}$$

其中 $c_1 = \sum_{k=0}^{q-1} \sum_{l \neq k} f(k,l)$, $c_2 = (q-1)\sum_{k=0}^{q-1} f(k,k)/c_1$.

引理 6.12 中的 $F(\cdot,\cdot)$ 是非限制性的, 可以被许多常用的偏差和其他测度满足. 由 (6.14) 式可得 $c_2 > 1$, 从而(6.15)式中 $A_i(\mathcal{P})$ 的系数随着 i 的增大而呈指数减小. 于是, 当考虑所有水平置换时, $\bar{\phi}(\mathcal{P})$ 的趋向与 GMA 是一致的. Tang 等 (2012) 与 Tang 和 Xu (2013) 也说明了当考虑所有水平置换时, CD 的均值是 GMP 的一个线性函数. 这里, 我们将他们的结果推广至一般的由再生核定义的偏差中.

注意到 (2.36) 式中的偏差可以表示成 (6.12) 式定义的空间填充测度如下:

$$F(\boldsymbol{x}_i, \boldsymbol{x}_j) = \mathcal{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) - \mathcal{K}_1(\boldsymbol{x}_i) - \mathcal{K}_1(\boldsymbol{x}_j) + \mathcal{K}_2, \tag{6.16}$$

其中 $\mathcal{K}_1(\boldsymbol{x}) = \int_{\mathcal{X}} \mathcal{K}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}F_u(\boldsymbol{y})$, $\mathcal{K}_2 = \int_{\mathcal{X}^2} \mathcal{K}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}F_u(\boldsymbol{x})\mathrm{d}F_u(\boldsymbol{y})$ 为常数. 换句话说, 任何由再生核定义的偏差都是空间填充测度的一个特例.

文献中常用的偏差的再生核是定义在 $\mathcal{X} = [0,1]^s$ 上的, 且有乘积形式:

$$\mathcal{K}(\boldsymbol{x},\boldsymbol{y}) = \prod_{k=1}^{s} f(x_k, y_k), \tag{6.17}$$

其中 $f(x,y)$ 是定义在 $[0,1]^2$ 上的. (2.36) 式中对应的偏差可表示为

$$D^2(\mathcal{P},\mathcal{K}) = \mathcal{K}_2 - \frac{2}{n}\sum_{i=1}^{n}\prod_{k=1}^{s} f_1(x_{ik}) + \frac{1}{n^2}\sum_{i,j=1}^{n}\prod_{k=1}^{s} f(x_{ik}, x_{jk}), \tag{6.18}$$

其中 $f_1(x) = \int_0^1 f(x,y)\mathrm{d}y$. 各种均匀性测度的核函数表达如下:

(i) 对 CD, $f(x,y) = 1 + (|x-0.5| + |y-0.5| - |x-y|)/2$;

(ii) 对 WD, $f(x,y) = 1.5 - |x-y| + |x-y|^2$;

(iii) 对 MD, $f(x,y) = 15/8 - |x-0.5|/4 - |y-0.5|/4 - 3|x-y|/4 + |x-y|^2/2$;

(iv) 对 Lee 偏差, $f(x,y) = 1 - \min\{|x-y|,\ 1-|x-y|\}$.

现在考虑任意给定的部分因子设计 $\mathcal{P}$ 的水平置换, 并计算所有置换设计的偏差均值, 记为 $\overline{D}(\mathcal{P},\mathcal{K})$. 对于 $\mathcal{P}$ 的任意一行 $\boldsymbol{x}_i$, 当考察所有的水平置换时, Z_q^s 的每个 s 元组均出现 $((q-1)!)^s$ 次, 于是

$$\sum_{\mathcal{P}'\in\mathcal{H}(\mathcal{P})}\sum_{i=1}^{n}\prod_{k=1}^{s}f_1(x_{ik})=\sum_{i=1}^{n}\sum_{\mathcal{P}'\in\mathcal{H}(\mathcal{P})}\prod_{k=1}^{s}f_1(x_{ik})$$
$$=n\left((q-1)!\sum_{k=0}^{q-1}f_1(k)\right)^s \tag{6.19}$$

为常数. 由引理 6.12, (6.18) 式, (6.19) 式与文献 (Zhou and Xu, 2014), 可以得到下面定理.

定理 6.13 假设 $\mathcal{K}(\boldsymbol{x},\boldsymbol{y})=\prod_{k=1}^{s}f(x_k,y_k)$ 且 $f(\cdot,\cdot)$ 满足 (6.14) 式. 对于一个 (n,q^s) 设计 $\mathcal{P}$, 当考虑 $\mathcal{P}$ 中所有水平置换时, 有

$$\overline{D}(\mathcal{P},\mathcal{K})=\mathcal{K}_0+\left(\frac{c_1(c_2+q-1)}{q^2(q-1)}\right)^s\sum_{i=0}^{s}\left(\frac{c_2-1}{c_2+q-1}\right)^iA_i(\mathcal{P}), \tag{6.20}$$

其中$\mathcal{K}_0=\mathcal{K}_2-2\left(\sum_{k=0}^{q-1}f_1((k+0.5)/q)/q\right)^s$ 为常数, $\mathcal{K}_2$ 和 $f_1(\cdot)$ 分别由 (6.16) 式和 (6.18) 式定义, $c_1=\sum_{k=0}^{q-1}\sum_{l\neq k}f((k+0.5)/q,(l+0.5)/q)$ 和 $c_2=(q-1)\sum_{k=0}^{q-1}f((k+0.5)/q,(k+0.5)/q)/c_1$.

由定理 6.13, 对于任何满足 (6.14) 式的由再生核定义的偏差, 并考虑所有水平置换时, 偏差的均值均是 GMP 的线性组合. 而且 WD, CD 等常用的偏差下均满足 (6.14) 式中的条件. 比如, 由 WD 的核函数, 有 $\mathcal{K}_2=(4/3)^s$, $\sum_{k=0}^{q-1}f_1((k+0.5)/q)/q=4/3$, $c_1=\sum_{k=0}^{q-1}\sum_{l\neq k}\left(1.5-|k-l|/q+|k-l|^2/q^2\right)=(q-1)(8q-1)/6$, $c_2=9q/(8q-1)>1$. 对于 WD 的均值, 有

$$\overline{\mathrm{WD}}(\mathcal{P})=-\left(\frac{4}{3}\right)^s+\left(\frac{8q^2+1}{6q^2}\right)^s\sum_{i=0}^{s}\left(\frac{q+1}{8q^2+1}\right)^iA_i(\mathcal{P}). \tag{6.21}$$

特别地, 当 $q=2$ 时, 所有的水平置换设计都有相同的 WD 值, 这是因为坐标旋转不改变 WD 值, 所以, (6.21) 式说明了二水平设计的 GMP 和 WD 之间的精确关系. 当 $q=3$ 时, Tang 等 (2012) 说明了当考虑所有水平置换时, 要想计算 $\bar{\phi}(\mathcal{P})$, 只需在线性水平置换下考虑即可. 然后再基于 (2.23) 式中 WD 的表达式可知, 任何线性水平置换并不改变 WD 值, 这就意味着 (6.21) 式也说明了三水平设计的 GMP 与 WD 之间精确的关系. 换句话说, (6.21) 式包含了 (Ma and Fang, 2001) 中 WD 下二水平和三水平设计的结果. 类似地, 将 CD 运用到定理 6.13 中, 可以得到 CD 和 GMP 之间的关系, 二水平、三水平和多水平设计的相关结论在 (Fang and Mukerjee, 2000; Ma and Fang, 2001; Tang et al., 2012; Tang and Xu, 2013) 中均有提及. 此外, 定理 6.13 包含了 (Zhou et al., 2008) 中二水平和三水平设计在 Lee 偏差下的相关结果, 以及 (Zhou et al., 2013) 中二水平设计在 MD 下的结果. 总的来说, 定理 6.13 给出了由再生核定义的任意均匀性测度的统一结果.

在这一节结束之前, 值得注意的是 Sun 等 (2011a) 研究了定理 6.6 和定理 6.9 中的密切联系. 这种研究是通过将 WD 表示为 (4.8) 式中的二次型, 以及把离散偏差表示为下面的二次型:

$$\mathrm{DD}^2(\mathcal{P}) = -\left[\frac{a+(q-1)b}{q}\right]^s + \frac{b^s}{n^2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}_s\boldsymbol{y},$$

其中 $\boldsymbol{y}$ 的定义与 (6.2) 式中一致, $\boldsymbol{D}_s = \otimes^s \boldsymbol{D}_0$, $\boldsymbol{D}_0 = (d_{ij})$, 当 $i = j$ 时, $d_{ij} = a/b$; 否则, $d_{ij} = 1$, $i, j = 0, \cdots, q-1$. 这种用示性函数的二次型来表达的偏差对寻找任意准则下的最优设计是很有帮助的. 事实上, 为了找到使得二次型最小的设计, 像 $\boldsymbol{y}^{\mathrm{T}}\mathsf{B}\boldsymbol{y}$, 只需解下列方程组:

$$\begin{cases} \min\limits_{\boldsymbol{y}} & \boldsymbol{y}^{\mathrm{T}}\mathsf{B}\boldsymbol{y}, \\ \text{s.t.} & \boldsymbol{y}^{\mathrm{T}}\mathbf{1} = n, \ \boldsymbol{y}_i = 0, \cdots, n. \end{cases} \tag{6.22}$$

Sun 等 (2009) 已经提供了解决此问题的一种方法, 许多在 GMA 和均匀性准则下的最优设计在这里也有列表. 对于更多解决像 (6.22) 式问题的方法, 可以参见 4.3 节.

6.5 投影均匀性及相关准则

已有文献中提出了许多准则来比较 U-型设计, 但这些准则均不能直接区分出非同构的饱和设计. 同构性的定义及相关讨论可参见 6.1 节. 设 $\mathcal{P}$ 是一个有 n 次试验、s 个因子的设计, 它有许多不同的子设计, 如 s 个包含一个因子的子设计, $\binom{s}{2}$ 个包含两因子的子设计, 以此类推. 某个特定的准则能够度量所有的子设计, 并称相应的数值为设计的投影型, 于是我们可以用这些投影值的分布或向量作为区分设计的工具. Hickernell 和 Liu (2002) 提出了*投影偏差型* (projection discrepancy pattern) 和相关准则, Fang 和 Qin (2004) 则建议对二水平因子设计使用一个不同的均匀型和相关的准则, 并基于优势理论和投影混杂提出了最小低阶混杂优势准则. 本节将针对这些方法给出简要的介绍.

6.5.1 投影偏差型及相关准则

本小节将介绍由 Hickernell 和 Liu (2002) 提出的*投影偏差型*及相关的准则.

每种偏差均可认为是在特定的核函数 (详见定义 2.34) 下的一种测度. 为定义设计的*投影偏差型*, 首先要限制*试验区域* $\mathcal{X}$ 为一维区域的笛卡儿积, 即 $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_s$. 假设*再生核* $\mathcal{K}$ 是一维再生核 $\mathcal{K}_j = 1 + \hat{\mathcal{K}}_j$ 的乘积, 即 $\mathcal{K} = \prod_{j=1}^s \mathcal{K}_j$. 又设 $\mathcal{S}$ 为集合 $\{1, \cdots, s\}$, u 为 $\mathcal{S}$ 的任一子集, $\boldsymbol{x}_u$ 为向量 $\boldsymbol{x}$ 中下标为子集 u 的元素, $\mathcal{X}_u$ 为 $j \in u$ 的 $\mathcal{X}_j$ 的笛卡儿积. 设 $\mathcal{P} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 为在区域 $\mathcal{X}$ 上的 n 个

试验点组成的设计, $\mathcal{P}_u$ 为设计 $\mathcal{P}$ 在区域 $\mathcal{X}_u$ 上的投影, 于是很自然地我们会希望当 $|u|$ 较小时投影 $\mathcal{P}_u$ 是 $\mathcal{X}_u$ 上好的设计. 分辨度 (Box et al., 1978, 385 页) 和混杂 (Fries and Hunter, 1980) 的概念正是基于此想法提出的. 前面定义的偏差并不能保证得到好的设计, 但是可以基于这些平方偏差重新定义一种混杂的概念. 现在回顾一下 Hickernell 和 Liu (2002) 提出的有关定义, 他们将再生核 $\mathcal{K}$ 表示为

$$\mathcal{K}(\boldsymbol{x},\boldsymbol{w})=\prod_{j=1}^{s}\mathcal{K}_j(x_j,w_j)=\sum_{\varnothing\subseteq u\subseteq\mathcal{S}}\hat{\mathcal{K}}_u(\boldsymbol{x}_u,\boldsymbol{w}_u), \tag{6.23a}$$

$$\hat{\mathcal{K}}_u(\boldsymbol{x}_u,\boldsymbol{w}_u)=\prod_{j\in u}\hat{\mathcal{K}}_j(x_j,w_j). \tag{6.23b}$$

注意到 $\hat{\mathcal{K}}_\varnothing=1$.

例 6.2 对 $j=1,\cdots,s$, 假设第 j 个因子有 q_j 个水平, 即 $\mathcal{X}_j=\{0,1,\cdots,q_j-1\}$, 并定义再生核为

$$\mathcal{K}_j(x,w)=1+\gamma(-1+q_j\delta_{xw}).$$

于是有

$$\hat{\mathcal{K}}_u(\boldsymbol{x}_u,\boldsymbol{w}_u)=\gamma^{|u|}\prod_{j\in u}(-1+q_j\delta_{x_jw_j}).$$

若取核函数如 (6.23) 式所示, 则

$$D^2(\mathcal{P};\mathcal{K})=\sum_{\varnothing\subset u\subseteq\mathcal{S}}D^2(\mathcal{P}_u;\hat{\mathcal{K}}_u)=\sum_{j=1}^{s}D^2_{(j)}(\mathcal{P};\mathcal{K}), \tag{6.24a}$$

其中

$$D^2_{(j)}(\mathcal{P};\mathcal{K})=\sum_{|u|=j}D^2(\mathcal{P}_u;\hat{\mathcal{K}}_u). \tag{6.24b}$$

又因为 $\hat{\mathcal{K}}_\varnothing=1$, 所以有 $D(\mathcal{P};\hat{\mathcal{K}}_\varnothing)=0$.

定义 6.4 假设 $\mathcal{X}=\mathcal{X}_1\times\cdots\times\mathcal{X}_s$, 且再生核 $\mathcal{K}$ 如 (6.23) 式所示, 则称 s 维向量

$$\mathrm{PD}(\mathcal{P};\mathcal{K})=(D_{(1)}(\mathcal{P};\mathcal{K}),\cdots,D_{(s)}(\mathcal{P};\mathcal{K}))$$

为投影偏差型(PDP). 字典序可以用来比较任两个设计 $\mathcal{P},\tilde{\mathcal{P}}\subseteq\mathcal{X}$: 称设计 $\mathcal{P}$ 比 $\tilde{\mathcal{P}}$ 有更好的投影均匀性, 等价地, 即 $\mathrm{PD}(\mathcal{P};\mathcal{K})<\mathrm{PD}(\tilde{\mathcal{P}};\mathcal{K})$ 当且仅当 $\mathrm{PD}(\mathcal{P};\mathcal{K})-\mathrm{PD}(\tilde{\mathcal{P}};\mathcal{K})$ 的第一个 (从左往右) 非零元是负数. 如果 $t=\min\{j:D_{(j)}(\mathcal{P};\mathcal{K})>0\}$, 则称 $\mathcal{P}$ 的均匀性分辨度 (uniformity resolution) 为 t. 进一步, 如果对于一个设计, 没有其他设计比它有更好的投影均匀性, 则称该设计有最小投影均匀性 (minimum projection uniformity, MPU).

对例 6.2, $D^2_{(j)}(\mathcal{P};\mathcal{K})$ 的表达式为

$$D^2_{(j)}(\mathcal{P};\mathcal{K}) = \frac{\gamma^j}{n^2}\sum_{|u|=j}\sum_{i,k=1}^{n}\prod_{l\in u}(-1+q_l\delta_{x_{il}x_{kl}}). \tag{6.25}$$

定义 6.4 并没有要求设计是正规的部分因子设计 (简写为 FFD), 甚至不要求每个因子的水平数都是有限的, 它只假设了试验区域是单因子区域的笛卡儿积以及再生核是一维核函数的乘积.

对于 (6.25) 式的特殊例子, 定义 6.4 退化为 (Xu and Wu, 2001) 中提出的 GMA, 即 Hickernell 和 Liu (2002) 给出的如下结论.

定理 6.14 如果 PDP 的元素如 (6.25) 式所示, 则由定义 6.4 给出的 MPU 等价于 Xu 和 Wu (2001) 定义的 GMA, 即

$$\gamma^{-j}D^2_{(j)}(\mathcal{P};\mathcal{K}) = A_j(\mathcal{P}), \quad 对 \quad j=1,\cdots,s.$$

进一步, 对于两水平设计, 这里定义的 MPU 与 Tang 和 Deng (1999) 提出的最小 G_2 混杂也是等价的.

下面讨论一下偏差定义中出现的参数 γ. 该参数的取值对于比较不同设计的 PDP 值是没有任何影响的, 因此在考虑 PDP 时取 γ 为 1. 在该条件下, Liu 等 (2006) 研究了 MPU, GMA, 最小矩混杂 (minimum moment aberration) (Xu, 2003) 和最小 χ^2 准则 (minimum χ^2 criterion) (Liu et al., 2006) 之间的关系. 这些关系为每一个准则都提供了强有力的统计解释, 他们还得到了一些一般化的最优性结果, 不但统一了几个已有结果, 而且有利于构造非对称的超饱和设计.

但是 γ 值对不同设计的偏差的比较却是有影响的. 回顾 (6.24) 式, 其中的平方偏差是 $D^2_{(j)}(\mathcal{P};\mathcal{K})$ 的和, 并注意到如果每个 $\hat{\mathcal{K}}_j$ 都包含参数 γ, 则 $D^2_{(j)}(\mathcal{P};\mathcal{K})$ 就会包含 γ^j 项. 于是参数 γ 的值越大, j 较大的 $D^2_{(j)}(\mathcal{P};\mathcal{K})$ 项就有相对更大的权重, 因此 $\mathcal{P}$ 的高维投影就会表现出更好的均匀性, 反之, 参数 γ 的值越小, $\mathcal{P}$ 的低维投影会表现出更好的均匀性. 所以可以得到如下定理, 即比较两个设计的混杂程度等价于比较它们在 γ 趋于零时的偏差, 具体解释可参见文献 (Hickernell and Liu, 2002), 其证明比较直接, 在这里省略.

定理 6.15 假设再生核如 (6.23) 式所示, 且 $\hat{\mathcal{K}}_j$ 包含一个系数 γ. 对于一个固定试验次数 n, 设 $\mathcal{P}_a$ 是 GMA 设计, t 是它的分辨度. 对任意的 $\gamma>0$, 设 $\mathcal{P}_\gamma$ 为对应的最小偏差设计, 那么如果 $D(\mathcal{P}_a;\mathcal{K})>0$, 则有

$$\lim_{\gamma\downarrow 0}\frac{D(\mathcal{P}_\gamma;\mathcal{K})}{D(\mathcal{P}_a;\mathcal{K})}=1, \quad \lim_{\gamma\downarrow 0}D_{(j)}(\mathcal{P}_\gamma;\mathcal{K}) = D_{(j)}(\mathcal{P}_a;\mathcal{K}) \quad (j=1,\cdots,t).$$

6.5.2 均匀性模式及相关准则

对于两水平 U-型设计, Fang 和 Qin (2004) 提出了均匀性模式及相关的准则. 设 $\mathcal{P}$ 为一个 $U(n;2^s)$ 设计, 为简单起见, 仅考虑 CD 和它对应的均匀性模式. 设 u 为 $\mathcal{S}$ 的任一子集, $\mathcal{P}_u$ 为 $\mathcal{P}$ 的子设计, 其 CD 写为 $\mathrm{CD}_{2,u}(\mathcal{P})$. 容易看到

$$[\mathrm{CD}_{2,u}(\mathcal{P})]^2=\left(\frac{1}{12}\right)^{|u|}-\frac{3^{|u|}}{2^{5|u|-1}}+\frac{1}{n2^{2|u|}}E_0(\mathcal{P}_u), \tag{6.26}$$

其中 $E_0(\mathcal{P}_u)$ 是子设计 $\mathcal{P}_u$ 的距离分布的第一个元素 (详见 (6.1)). 设

$$[I_i(\mathcal{P})]^2=\sum_{|u|=i}[\mathrm{CD}_{2,u}(\mathcal{P})]^2,$$

它是 $\mathcal{P}$ 在 i 子维上的整体均匀性的一种度量. 很显然, $\mathrm{CD}(\mathcal{P})$ 可分解为 $\mathrm{CD}^2(\mathcal{P})=\sum_{i=1}^{s}[I_i(\mathcal{P})]^2$. 从 (6.26) 式中可以明显看到 $\mathcal{P}$ 的强度为 t 当且仅当对于 $1\leqslant j\leqslant t$, 有

$$[I_j(\mathcal{P})]^2=\binom{s}{j}\left[\left(\frac{1}{12}\right)^{j}-\frac{3^{j}}{2^{5j-1}}+\frac{1}{8^{j}}\right],$$

且

$$[I_{t+1}(\mathcal{P})]^2>\binom{s}{t+1}\left[\left(\frac{1}{12}\right)^{t+1}-\frac{3^{t+1}}{2^{5(t+1)-1}}+\frac{1}{8^{(t+1)}}\right].$$

定义 6.5 定义

$$U_i(\mathcal{P})=[I_i(\mathcal{P})]^2-\binom{s}{i}\left[\left(\frac{1}{12}\right)^{i}-\frac{3^{i}}{2^{5i-1}}+\frac{1}{8^{i}}\right],\quad 1\leqslant i\leqslant s. \tag{6.27}$$

则称向量 $(U_1(\mathcal{P}),\cdots,U_s(\mathcal{P}))$ 为设计 $\mathcal{P}$ 的均匀性模式 (uniformity pattern, UP).

接下来, 基于该模式的 MPU 准则就可以用类似定义 6.4 的方式得到.

文献 (Fang and Qin, 2004) 中给出了以下定理, 其中第二个定理建立了 $\{U_i(\mathcal{P})\}$ 和 $\{A_i(\mathcal{P})\}$ 之间的解析关系.

定理 6.16 假设 $\mathcal{P}\in\mathcal{U}(n;2^s)$, 则 $\mathcal{P}$ 的强度为 t 当且仅当 $U_j(\mathcal{P})=0, 1\leqslant j\leqslant t$, 且 $U_{t+1}(\mathcal{P})\neq 0$.

定理 6.17 假设 $\mathcal{P}\in\mathcal{P}(n;2^s)$, 则对任意 $1\leqslant i\leqslant s$, U_j 和 A_j 可以通过下面的式子相互线性表示

$$U_i(\mathcal{P})=\frac{1}{8^i}\sum_{v=1}^{i}\binom{s-v}{s-i}A_v(\mathcal{P}) \tag{6.28}$$

和

$$A_i(\mathcal{P})=\sum_{v=1}^{i}(-1)^{i-v}\binom{s-v}{i-v}8^vU_v(\mathcal{P}). \tag{6.29}$$

需要注意的是, (6.28) 式中系数 $\frac{1}{8^i}\binom{s-v}{s-i}$ 是正数. 显然, 对于 $i=1,\cdots,s$ 依次最小化 $U_i(\mathcal{P})$ 等价于依次最小化 $A_i(\mathcal{P})$. 因此, 对于二水平设计, MPU 准则和 GMA 准则是一致的, 特别地, 一个二水平设计有 GMA 当且仅当它有 MPU.

MPU 与其他准则之间的关系还有许多地方需要我们研究. 对于二水平设计, Zhang 和 Qin (2006) 说明了 MPU 准则、Tang (2001) 提出的 V-准则与 Fang 等 (2003b) 提出的近似正交性准则是相互等价的, 且常用的 $E(s^2)$-最优就是 MPU 的一个特例. Qin 等 (2011) 还发现 MPU 准则是 Cheng 等 (2002) 提出的设计有效性准则的一个很好的替代. 而对于多水平设计, Qin 等 (2012) 讨论了基于 Qin 和 Fang (2004) 提出的离散偏差得到的 MPU 准则的相关问题, 并给出了 MPU 与其他最优性准则之间的联系.

6.6 优势理论框架

令 $\boldsymbol{U}$ 是 $\mathcal{U}$ 上的一个 $n\times s$ 设计. 在 $\mathcal{U}$ 上评估设计好坏的准则可以使用函数 $\phi(\boldsymbol{U})$ 或者一个向量 $(\phi_1(\boldsymbol{U}),\cdots,\phi_m(\boldsymbol{U}))$. 字长型就是一个使用向量作为准则的例子. 怎么比较两个这样的向量? 1.4.1 小节介绍了两种向量的排序方法: 字典顺序和优势顺序. 本节关心的是两个不同的向量函数: 两两相遇数和优势顺序下的广义字长型.

6.6.1 基于两两相遇数向量

令 $\delta_{ik}(\boldsymbol{U})$ 是设计矩阵 $\boldsymbol{U}=(u_{ik})$ 的 $\boldsymbol{u}_i$ 和 $\boldsymbol{u}_k$ 两行之间的**两两相遇数** (PC). 显然, $\boldsymbol{u}_i$ 和 $\boldsymbol{u}_k$ 行之间的相遇数 (参见 (2.48)) 为 $\delta_{ik}(\boldsymbol{U})=s-d_{ik}^H(\boldsymbol{U})$, 其中 $d_{ik}^H(\boldsymbol{U})$ 是第 $\boldsymbol{u}_i$ 行和第 $\boldsymbol{u}_k$ 两行之间的 Hamming 距离. 对 $1\leqslant i<k\leqslant n$, 令 $\boldsymbol{\delta}(\boldsymbol{U})=(\delta_1,\delta_2,\cdots,\delta_m)^{\mathrm{T}}$ 为 $\delta_{ik}(\boldsymbol{U})$ 的**两两相遇数向量** (PC 向量), 这里 δ_{ik} 对应于元素 $\delta_{n(i-1)+k-i(i+1)/2}$, 其中 $m=n(n-1)/2$, 即

$$\boldsymbol{\delta}(\boldsymbol{U})=(\delta_{12}(\boldsymbol{U}),\cdots,\delta_{1n}(\boldsymbol{U}),\delta_{23}(\boldsymbol{U}),\cdots,\delta_{2n}(\boldsymbol{U}),\cdots,\delta_{(n-1)n}(\boldsymbol{U})). \tag{6.30}$$

如果两个设计的 PC 向量不能通过元素置换得到, 那么就说它们的PC *不同*. 对任意的 i,j, 注意到 $1+\sum_{k\neq i}\delta_{u_{ij}u_{kj}}=n/q$, 可以发现 PC 和

$$\sum_{r=1}^{m}\delta_r=\sum_{i<k}\sum_{j=1}^{s}\delta_{u_{ij}u_{kj}}$$

对任何的 U-型设计均是保持不变的.

引理 6.18 对任何 $\boldsymbol{U}\in\mathcal{U}(n,q^s)$, 它的 PC 和是 $\frac{ns}{2}\left(\frac{n}{q}-1\right)$.

基于决策论、优势理论和引理 6.18, 利用 PC 向量, Zhang 等 (2005) 对 U-型设计提出了一个一般框架.

如果选择某个合适的 Schur- 凸函数 $\Psi(\boldsymbol{\delta}(\boldsymbol{U}))$ ($\Psi(\boldsymbol{x}) = \sum_{r=1}^{m} \psi(x_r)$, 其中 ψ 在 R_+ 上是凸的, 请参见 1.4.1 小节), 则函数 $\Psi(\boldsymbol{\delta}(\boldsymbol{U}))$ 可以作为比较 $\mathcal{U}(n, q^s)$ 上两个 U-型设计的一个准则.

定义 6.6 在设计空间 $\mathcal{U} = \mathcal{U}(n, q^s)$ 上定义:

1) 可容许性: 如果 $\boldsymbol{\delta}(\boldsymbol{U}_1) \prec \boldsymbol{\delta}(\boldsymbol{U}_2)$, 就说设计 $\boldsymbol{U}_1$ 比 $\boldsymbol{U}_2$ 更好. 如果存在一个设计 $\boldsymbol{U}^*$, 使得 $\boldsymbol{\delta}(\boldsymbol{U}^*) \prec \boldsymbol{\delta}(\boldsymbol{U})$, 那么就说设计 $\boldsymbol{U}$ 是不可容许的. 一个设计如果不是不可容许的, 那么就称该设计是容许的.

2) 优势设计: 如果存在一个设计 $\boldsymbol{U}$ 使得

$$\boldsymbol{\delta}(\boldsymbol{U}) \preceq \boldsymbol{\delta}(\boldsymbol{U}^*), \quad \text{对任何} \quad \boldsymbol{U}^* \in \mathcal{U},$$

就称 $\boldsymbol{U}$ 是空间 $\mathcal{U}$ 上的一个优势设计.

3) Schur-凸最优性: 对先前 R_+^m 上定义的 Schur-凸核函数 $\Psi(\cdot)$, 设计 $\boldsymbol{U}$ 称为关于 $\Psi(\cdot)$ Schur-Ψ-最优的, 如果

$$\text{对任何} \quad \boldsymbol{U}^* \in \mathcal{U}, \quad \text{都有} \quad \Psi(\boldsymbol{\delta}(\boldsymbol{U})) \leqslant \Psi(\boldsymbol{\delta}(\boldsymbol{U}^*)).$$

上面定义中的三部分可以分层次地分为两个阶段的研究, 即严密的优势检查和灵活的 Schur-凸比较. 在第一个阶段, 对 $\mathcal{U}$ 上的不同设计, 计算它们元素按增序排列的 PC 向量并比较优势顺序 (1.32) 的累积和. 根据定义 6.6, 任何不容许的设计应该在试验中被禁止; 如果优势设计存在, 那它就是最优的, 应该被推荐. 否则, 转到第二阶段去比较可容许设计. 第一阶段是严密的, 因为优势比较需要 PC 向量间严格的条件. 在第二阶段, 指定一个凸的核函数并且计算每一个可容许设计的 Schur-Ψ 值. 因为上述的 Schur-Ψ 准则是单值的, 故而所有的设计可以两两比较且是可以排序的. 对不同的目的, 可以灵活地预先定义核函数, 只要它们是凸函数即可. 上面两个阶段是基于如下理论.

定理 6.19 对于两个设计 $\boldsymbol{U}_1, \boldsymbol{U}_2 \in \mathcal{U}(n, q^s)$, $\boldsymbol{\delta}(\boldsymbol{U}_1) \preceq \boldsymbol{\delta}(\boldsymbol{U}_2)$ 当且仅当 $\Psi(\boldsymbol{\delta}(\boldsymbol{U}_1)) \leqslant \Psi(\boldsymbol{\delta}(\boldsymbol{U}_2))$, 其中 $\Psi(\cdot)$ 是任意一个 Schur-凸函数.

Zhang 等 (2005) 用一个例子和一些凸函数解释了上述观点. 他们的例子说明了优势框架用来评估设计的严密性和灵活性. 形式上, 他们有如下的定理去刻画优势设计和 Schur-Ψ 优势设计间的充要条件.

Zhang 等 (2005) 发现现有的很多准则, 包括离散偏差, $q = 2, 3$ 情形下的 WD ($q = 2, 3$) 情形下的 CD, 都可以被表示为一个可分离的凸函数, 即形式为 (1.33) 的 Schur-Ψ 函数. 该事实可以扩展到更多的准则, 例如, 超饱和设计 (参见第 7 章) 中

用到的 $E(s^2)$, $\mathrm{Ave}(\chi^2)$ 和 $\mathrm{Ave}(f^2)$, 因此 Zhang 等 (2005) 提出了一个寻找可分离凸函数紧下界的统一方法.

考虑 U-型设计 $\boldsymbol{U} \in \mathcal{U}(n; q^s)$ 的 PC 均值, 根据引理 6.18, PC 均值是常数 $\bar{\delta} = \dfrac{s(n-q)}{q(n-1)}$. 现在可把引理 1.1 和引理 1.2 用在 $\mathcal{U}$ 上的 $\boldsymbol{\delta}(\boldsymbol{U})$ 向量集. 对长度为 m 的整值 $\boldsymbol{\delta}(\boldsymbol{U})$, 令

$$\overline{\boldsymbol{\delta}} \equiv (\underbrace{\bar{\delta}, \cdots, \bar{\delta}}_{m})^{\mathrm{T}} \quad 和 \quad \widetilde{\boldsymbol{\delta}} \equiv (\underbrace{\theta, \cdots, \theta}_{m(1-f)}, \underbrace{\theta+1, \cdots, \theta+1}_{mf})^{\mathrm{T}},$$

其中 θ 和 f 分别是 $\bar{\delta}$ 的整数部分和分数部分. 显然有 $\overline{\boldsymbol{\delta}} \preceq \widetilde{\boldsymbol{\delta}} \preceq \boldsymbol{\delta}$, 其中当 $f=0$ 时, $\widetilde{\boldsymbol{\delta}}$ 可简化为 $\overline{\boldsymbol{\delta}}$.

定理 6.20 一个 U-型设计是优势设计当且仅当它对每一个凸核函数是 Schur-最优的. 对任一明确定义的 Schur-ψ 准则, 均有下界 $m(1-f)\psi(\theta) + mf\psi(\theta+1)$.

(Dey and Mukerjee, 1999) 中的引理 5.2.1 在上述定理的证明中有重要的作用. Zhang 等 (2005) 也讨论了在考察设计的正交性、混杂和均匀性时, 怎样选择合适的核函数, 同时也提出了一些新的准则.

6.6.2 最小混杂优势

在 6.6.1 小节中基于 PC 向量的优势框架显示了它在比较 U-型设计中的完美表现. 然而, 任一好的准则在某些情形下会遇到困题. 例如, 一个饱和正交表 $L_n(q^s)$ 的任两行的 Hamming 距离均是 n/q (参见 (Mukerjee and Wu, 1995) 与定义 1.4). 因此, PC 向量和优势结构不能区别非同构的饱和正交设计, 这是因为它们有相同的 PC 向量. 如果一个设计可以通过另一个设计重新标记因子, 重新安排试验顺序或者交换因子的水平得到, 那么它们就被称为*同构的*. 一个著名的例子是关于 $L_{16}(2^{15})$, 它有五个非同构的 $L_{16}(2^{15})$ 设计, 根据 Hall (1961) 给的顺序, 记为 HM16.1 到 HM16.5. 我们在下面列出前两个设计 HM16.1 和 HM16.2, 其中用 0 和 1 来表示两个水平. 有些作者对如何区分它们和它们的子设计感兴趣. Sun 和 Wu (1993) 通过不完全定义对照子群定义了字长型, 并研究这五个设计的代数结构. Lin 和 Draper (1995) 考虑了这五个设计到不多于 5 维上的纯几何投影模式. Fang 和 Zhang (2004) 为了区分这几个设计提出了一个所谓的*最小混杂优势准则*. 对给定 $L_{16}(2^{15})$ 设计的每一个子设计 $L_{16}(2^p)(p \leqslant 15)$, 可以计算它的广义字长型 (GWP) (参见 1.4 节). 对一个给定的 p 共有 $N_p = \dbinom{s}{p}$ 个子设计, 一个很自然的想法是考虑所有 N_p 个子设计的平均投影 GWP. Fang 和 Zhang (2004) 发现非同构饱和设计的平均投影 GWP 是一致的, 且平均投影 GWP 可以由整体 GWP 表示, 即 $L_n(q^s)$ 中所有设计 (不饱

和或者饱和) 只要有同样的整体 GWP, 那么它们的平均投影 GWP 也是一样的.

$$\mathrm{HM16.1}=\begin{pmatrix}
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
0&1&0&1&0&1&0&1&0&1&0&1&0&1&0\\
1&0&0&1&1&0&0&1&1&0&0&1&1&0&0\\
0&0&1&1&0&0&1&1&0&0&1&1&0&0&1\\
1&1&1&0&0&0&0&1&1&1&1&0&0&0&0\\
0&1&0&0&1&0&1&1&0&1&0&0&1&0&1\\
1&0&0&0&0&1&1&1&1&0&0&0&0&1&1\\
0&0&1&0&1&1&0&1&0&0&1&0&1&1&0\\
1&1&1&1&1&1&1&0&0&0&0&0&0&0&0\\
0&1&0&1&0&1&0&0&1&0&1&0&1&0&1\\
1&0&0&1&1&0&0&0&0&1&1&0&0&1&1\\
0&0&1&1&0&0&1&0&1&1&0&0&1&1&0\\
1&1&1&0&0&0&0&0&0&0&0&1&1&1&1\\
0&1&0&0&1&0&1&0&1&0&1&1&0&1&0\\
1&0&0&0&0&1&1&0&0&1&1&1&1&0&0\\
0&0&1&0&1&1&0&0&1&1&0&1&0&0&1
\end{pmatrix}$$

和

$$\mathrm{HM16.2}=\begin{pmatrix}
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
0&1&0&1&0&1&0&1&0&1&0&1&0&1&0\\
1&0&0&1&1&0&0&1&1&0&0&1&1&0&0\\
0&0&1&1&0&0&1&1&0&0&1&1&0&0&1\\
1&1&1&0&0&0&0&1&1&1&1&0&0&0&0\\
0&1&0&0&1&0&1&1&0&1&0&0&1&0&1\\
1&0&0&0&0&1&1&1&1&0&0&0&0&1&1\\
0&0&1&0&1&1&0&1&0&0&1&0&1&1&0\\
1&1&1&1&1&1&1&0&0&0&0&0&0&0&0\\
0&1&0&1&0&0&1&0&1&0&1&0&1&1&0\\
1&0&0&1&1&0&0&0&0&1&1&0&0&1&1\\
0&0&1&1&0&1&0&0&1&1&0&0&1&0&1\\
1&1&1&0&0&0&0&0&0&0&0&1&1&1&1\\
0&1&0&0&1&1&0&0&1&0&1&1&0&0&1\\
1&0&0&0&0&1&1&0&0&1&1&1&1&0&0\\
0&0&1&0&1&0&1&0&1&1&0&1&0&1&0
\end{pmatrix}.$$

让我们用 $L_{16}(2^{15})$ 中的HM16.2 举例, 并取 $p=5$ 来解释投影 GWP 的分布. 总共有 $3003=\binom{15}{5}$ 个子设计, 9 种不同的 GWP: 1056 个子设计的 GWP 是 $(0,0,1,0,0)$, 384 个子设计的 GWP 为 $(0,0,0,1,0)$ 等. 实际上, HM16.2 的平均投影 GWP 与 HM16.$i(i=1,3,4,5)$ 是一致的. 通过将 GWP 和相应的频率写在一起, 将 HM16.2 的 3003 个 5 维投影 GWP 按 1.4.1 小节的字典顺序 $\models$ 分类如下:

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1056 \\ 0 & 0 & 0 & 1 & 0 & 384 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 576 \\ 0 & 0 & \frac{3}{2} & \frac{1}{2} & 0 & 288 \\ 0 & 0 & 0 & 0 & 1 & 72 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 192 \\ 0 & 0 & 2 & 1 & 0 & 99 \\ 0 & 0 & 1 & 1 & 0 & 144 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} & 192 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 72 \\ 0 & 0 & 0 & 1 & 0 & 384 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 192 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 576 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} & 192 \\ 0 & 0 & 1 & 0 & 0 & 1056 \\ 0 & 0 & 1 & 1 & 0 & 144 \\ 0 & 0 & \frac{3}{2} & \frac{1}{2} & 0 & 288 \\ 0 & 0 & 2 & 1 & 0 & 99 \end{pmatrix}.$$

现在可以定义新的概念 “投影 GWP 分布”如下.

定义 6.7　对一个饱和设计 $L_n(q^s)$, 它的 p 维 $(1 \leqslant p \leqslant s)$ 投影 GWP 的分布定义为它的 $N_p=\binom{n}{p}$ 个投影设计的广义字长型的分布. 我们符号化地把它写为一个顺序统计量 $\Psi_p=\left\{W_{[i]}\right\}_{i=1}^{N_p}$, 其中当 $i \leqslant j$ 时, $W_{[i]} \models W_{[j]}$.

令 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 是两个非同构的饱和 $L_n(q^s)$ 设计. Fang 和 Zhang (2004) 发现了

$$\sum_{i=1}^{N_p} W_{[i]}(\mathcal{P}_1)=\sum_{i=1}^{N_p} W_{[i]}(\mathcal{P}_2). \tag{6.31}$$

该事实为我们运用优势理论去比较非同构饱和设计提供了可能. 也可以类似地定义 $\Psi_p(\mathcal{P}_1)$ 和 $\Psi_p(\mathcal{P}_2)$ 间的优势关系.

定义 6.8 (最小混杂优势)　对于两个饱和的设计 $\mathcal{P}_1$ 和 $\mathcal{P}_2 \in L_n(q^s)$, p $(1 \leqslant p \leqslant s)$, 说 $\mathcal{P}_1$ 比 $\mathcal{P}_2$ 有更小的混杂优势且记为

$$\Psi_p(\mathcal{P}_1) \overset{\alpha}{\succeq} \Psi_p(\mathcal{P}_2) \quad (\text{或者当}\Psi_p(\mathcal{P}_1) \neq \Psi_p(\mathcal{P}_2)\text{时}, \Psi_p(\mathcal{P}_1) \overset{\alpha}{\succ} \Psi_p(\mathcal{P}_2)) \tag{6.32}$$

当且仅当

$$\sum_{i=1}^{t} W_{[i]}(\mathcal{P}_1) \models \sum_{i=1}^{t} W_{[i]}(\mathcal{P}_2), \quad t=1,2,\cdots,N_p-1. \tag{6.33}$$

一个饱和设计 $\mathcal{P}_*$ 被称为最小混杂优势(MAM), 如果没有设计比它有更小的混杂优势. 更近一步地, 如果对所有的 $p=1,\cdots,s$, $\Psi_p(\mathcal{P}_1)\overset{\alpha}{\succeq}\Psi_p(\mathcal{P}_2)$ 中至少有一个严格的优势关系 $\overset{\alpha}{\succ}$ 成立, 就说在 MAM 准则下, $\mathcal{P}_1$ 全局控制 $\mathcal{P}_2$, 记为 $\mathcal{P}_1\overset{\alpha}{\gg}\mathcal{P}_2$.

MAM 准则关注的是所有子设计中具有较小混杂能力的设计. 如果 $\mathcal{P}_1\overset{\alpha}{\gg}\mathcal{P}_2$, 则对任何 $m\ (1\leqslant m\leqslant N_p)$ 个 $p\ (p=1,\cdots,s)$ 维子设计, 可以找到 $\mathcal{P}_1$ 的 m 个子设计, 使得这 m 个子设计的平均 GWP 比 $\mathcal{P}_2$ 的任何 m 个子设计的平均 GWP 有更小的广义混杂. MAM 准则可以用来检测设计的非同构. 对 $L_n(q^s)$ 中的两个设计 $\mathcal{P}_1$ 和 $\mathcal{P}_2$, 如果它们是同构的, 那么它们有相同的 $\Psi_p\ (p=1,\cdots,s)$. 如果存在 p 使得 $\Psi_p(\mathcal{P}_1)\overset{\alpha}{\succ}\Psi_p(\mathcal{P}_2)$, 则表明它们是非同构的. 通过对 $L_{16}(2^{15})$ 的五个非同构设计应用 MAM 准则, Fang 和 Zhang (2004) 发现了

$$\text{HM16.1}\overset{\alpha}{\gg}\text{HM16.2}\overset{\alpha}{\gg}\text{HM16.3}\overset{\alpha}{\gg}\begin{cases}\text{HM16.4},\\ \text{HM16.5}.\end{cases}$$

还得到 HM16.4 和 HM16.5 在低维 $(p=1,\cdots,5)$ 和高维 $(p=10,\cdots,15)$ 上有同样的投影 GWP 分布, 但在维度为 $p=6,7,8,9$ 的投影 GWP 分布是不同的. HM16.4 和 HM16.5 并没有明显的优势关系.

习　题

6.1　证明下面两个设计 D_1 和 D_2 是同构的.

$$D_1=\begin{pmatrix}1&1&1\\1&-1&1\\1&1&-1\\-1&-1&-1\\-1&1&-1\\-1&-1&1\end{pmatrix},\quad D_2=\begin{pmatrix}-1&-1&-1\\1&-1&-1\\-1&1&-1\\1&1&1\\-1&1&1\\1&-1&1\end{pmatrix}.$$

并说明这两个设计是属于:

(i) U-型设计; (ii) 正交设计; (iii) 部分因子设计.

6.2　用 $\mathcal{U}_9(3^4)$ 表示水平为 1, 2 和 3 的所有可能的正交设计 $L_9(3^4)$ 的集合. 回答下面的问题:

1) 说明 $\mathcal{U}_9(3^4)$ 中的设计均是同构的.

2) 计算 $\mathcal{U}_9(3^4)$ 中所有设计的 WD, CD 和 MD, 并给出你的结论.

3) 表 6.4 列出了两个 $L_9(3^4)$ 设计, 其中 $\boldsymbol{U}_2$ 是 $\mathcal{U}_9(3^4)$ 上最小化 CD 得到的. 称 $\boldsymbol{U}_2$ 为 CD 意义下的均匀正交设计. 请找出 MD/WD 意义下的均匀正交设计.

4) 计算 $\mathcal{U}_9(3^4)$ 中所有设计的离散偏差值, 并给出你的结论.

5) 对均匀正交设计和均匀最小低阶混杂设计两个概念进行讨论.

6) 计算投影偏差型 (见定义 6.4).

6.3　计算表 1.2 中 $L_8(2^7)$ 的所有子设计 $L_8(2^5)$ 的均匀性模式 (见定义 6.5).

6.4　表 6.5 给了一个 MD 意义下的均匀最小低阶混杂设计 $\mathrm{UL}_{27}(3^{13})$, 该设计包含了 MD 意义下的几个均匀最小低阶混杂子设计 $\mathrm{UL}_{27}(3^s)(s<13)$. 请通过详细计算分别给出当 $s=6$ 和 $s=10$ 时这样的子设计.

6.5　共有五个非同构的 $L_{16}(2^{15})$ 设计. 本章已经列出了其中的两个, 请从文献中给出其他的三个.

6.6　共有两个非同构的 $L_{27}(3^{13})$ 设计 (参见文献 (Fang and Zhang, 2004)). 用算法 6.2 来检测它们的非同构性.

6.7　表 6.2 列出了四个正交设计 $L_{18}(3^7)$. Ma 和 Fang (2001) 指出至少有 3 个非同构的 $L_{18}(3^7)$. 更进一步地, Evangelaras 等 (2007) 证实了存在 3 个非同构的, 并且得到了几个最小低阶混杂子设计 $L_{18}(3^s)(s\leqslant 7)$. 请尽可能多地列出这些最小低阶混杂子设计.

第 7 章　均匀性在其他类型设计中的应用

从第 6 章可以看出, 均匀性和其他设计准则之间存在紧密的联系. 在本章中, 我们将说明均匀性准则在某些经典设计中也是一个有用的准则, 这些设计包括区组设计、超饱和设计、拉丁超立方体设计等.

7.1　区组设计中的均匀性

区组设计是一类重要的试验设计. 它的基本想法来源于农业和生物试验, 现在已经被广泛应用于科学和工程的很多领域. 最广泛使用的是**平衡不完全区组设计** (balanced incomplete block design, BIBD), 见定义 3.5. 另一个重要的设计是**可分解不完全区组设计** (resolvable incomplete block design, RIBD). 区组设计的基本概念可参见 3.6.1 小节. 而关于区组设计更详细的讨论, 可参见 (Dey, 1986).

我们知道, 区组设计的定义反映了一种区组间或处理间或组群间的"**平衡性**". 这种平衡性在直观上很易接受. 但在现有关于区组设计工作中, 平衡性的准则是从估计的角度引入的. 实际上, 平衡性准则可以看作是一种均匀性. Liu 和 Chan (2004) 与 Liu 和 Fang (2005) 研究了区组设计的均匀性并得到一些满意的结果.

7.1.1　BIBD 中的均匀性

BIBD 的特点在于将设计点均衡地排列. Liu 和 Chan (2004) 从均匀性角度研究了不完全区组设计. 他们利用 2.5.1 节定义的离散偏差理论上证明了在等重复、相宜的、二元不完全区组设计中, BIBD 是最均匀的.

假定将 n 个处理安排到 s 个区组中, 要求第 j 个区组有 t_j 个试验单元, 第 i 个处理在整个设计中出现 m_i 次, 这里 $i = 1, \cdots, n; j = 1, \cdots, s$. 令 z_i^j 为第 i 个处理出现在第 j 个区组中的次数, 则 $n \times s$ 矩阵 $\boldsymbol{Z} = (z_i^j)$ 称作该区组设计的**关联矩阵**. 根据定义 3.4, 记一个等重复 (即 $r_i = r$)、相宜的 (即 $t_j = t$) 和二元的 (即 $z_i^j = 1$ 或 0) 不完全区组设计为 $\mathrm{IB}(n, s, r, t)$. 把 s 个区组看作 s 个有 0 和 1 两个水平的因子, 把每个处理在 s 个区组上的分配看作一个取元素为 0 和 1 的点, 其中 1 表示该处理出现在相应的区组中, 0 则表示没有出现. 所以一个 $\mathrm{IB}(n, s, r, t)$ 设计的 n 个点正好对应关联矩阵 $\boldsymbol{Z}$ 的 n 个行. 对于 (2.49) 和 (2.50) 中定义的再生核, Liu 和 Chan (2004) 给出 $\boldsymbol{Z}$ 的离散偏差的平方为

$$\mathrm{DD}^2(\boldsymbol{Z}) = -\left(\frac{a+b}{2}\right)^s + \frac{a^s}{n} + \frac{a^s}{n^2}\left(\frac{b}{a}\right)^{2r}\sum_{i,j=1,j\neq i}^{n}\left(\frac{a}{b}\right)^{2n_{ij}^{11}}, \tag{7.1}$$

其中 n_{ij}^{11} 为同时包含处理 i 和 j 的区组的个数. 基于 (7.1), Liu 和 Chan (2004) 进一步得到如下定理.

定理 7.1 令 $\boldsymbol{Z}$ 为一个 IB(n,s,r,t) 设计的关联矩阵, 则

$$\mathrm{DD}^2(\boldsymbol{Z}) \geqslant -\left(\frac{a+b}{2}\right)^s + \frac{a^s}{n} + \frac{n-1}{n}a^s\left(\frac{b}{a}\right)^{2r-2\lambda}, \tag{7.2}$$

其中 $\lambda=r(t-1)/(n-1)$, 并且 (7.2) 式右边 $\mathrm{DD}^2(\boldsymbol{Z})$ 的下界能够达到的充要条件为 λ 是一个正整数且任一处理对都出现在 λ 个区组中, 即该设计是一个 BIBD(n,s,r,t,λ).

定理 7.1 表明, 对于一个 BIBD(n,s,r,t,λ), 假如它达到 $\mathrm{DD}^2(\boldsymbol{Z})$ 的下界, 则该 BIBD(n,s,r,t,λ) 是所有 IB(n,s,r,t) 设计中最均匀的. 该重要特征正是 BIB 设计的均衡性.

7.1.2 PRIBD 中的均匀性

此外, Liu 和 Fang (2005) 考虑了等重复的二元不完全区组设计, 不关心其相宜与否. 若一个区组设计是可分解的, 则它也是等重复和二元的. 令 $\boldsymbol{t}=(t_1,t_2,\cdots,t_s)^{\mathrm{T}}$, 用 RIB$(n,s,m,\boldsymbol{t},\boldsymbol{Z})$ 表示一个可分解的不完全区组设计 (简称为 RIB 设计). 特别地, 对一个 RIB$(n,s,m,\boldsymbol{t},\boldsymbol{Z})$, 如果每一个平行类是相宜的, 即在第 i 个 $(1\leqslant i\leqslant m)$ 平行类中, 有 q_i 个大小为 n/q_i 的区组, 则称该设计为一个 PRIB 设计, 记作 PRIBD$(n,s,m,\boldsymbol{t},\boldsymbol{Z})$. 这里 $q_1,\cdots,q_m$ 是 n 的正约数. Liu 和 Fang (2005) 给出了离散偏差在 (2.54) 下一个 PRIBD 是最均匀的充分必要条件.

定理 7.2 对一个 PRIBD$(n,s,r,\boldsymbol{t},\boldsymbol{Z})$, 有

$$\mathrm{DD}^2(\boldsymbol{Z}) \geqslant -\left(\frac{a+b}{2}\right)^s + \frac{a^s}{n} + \frac{n-1}{n}a^s\left(\frac{b}{a}\right)^{2r-2\lambda}, \tag{7.3}$$

其中 $\lambda=\left(\sum_{k=1}^r n/q_k - r\right)/(n-1)$, (7.3) 式右边 $\mathrm{DD}^2(\boldsymbol{Z})$ 的下界能够达到当且仅当 λ 是一个正整数且任一处理对都出现在 λ 个区组中.

基于该定理, 称一个 PRIBD$(n,s,m,\boldsymbol{t},\boldsymbol{Z})$ 是一个均匀 PRIB 设计, 如果其 $\mathrm{DD}^2(\boldsymbol{Z})$ 值达到 (7.3) 式中的下界. 在这样一均匀 PRIB 设计中, 任一处理对都出现在相同个数的区组中, 这意味着在处理之间存在着某种 "平衡". 实际上, 当一个均匀 PRIB 设计还是相宜的时, 即对所有 $1\leqslant i\leqslant m$, $q_i=q$(比方说), 那么由该定理容易得到下面的结论.

推论 7.3 如果一个均匀 PRIBD$(n,s,r,\boldsymbol{t},\lambda)$ 是相宜的, 则该设计是一个可分解的 BIBD$(n,rq,r,n/q,\lambda)$, 其中对所有 $1\leqslant i\leqslant s$, $t_i=n/q$, 且 $\lambda=r(n/q-1)/(n-1)$.

由该推论可知, “平衡” 可以看作是一种均匀性.

Liu 和 Fang (2005) 进一步说明对于一个均匀 PRIBD, 所有的基本处理对照都是可估的, 区组效应之间的基本对照也是可估的. 他们又给出一种通过一种 U-型设计构造均匀 PRIB 设计的方法. 这种方法在 U-型设计和 PRIBD 之间建立起一重要桥梁.

所有的结论证实了我们的判断, 即 “平衡” 准则可以看作是一种均匀性. 这些结果是在 (2.54) 中的离散偏差的意义下得到的, 但采用任何一个由 Hickernell (1998a, 1998b) 提出的反射不变的 L_2-偏差, 即中心化偏差 (CD)、可卷偏差 (WD)、对称化 L_2-偏差和未锚 L_2-偏差, 结论同样成立. 前两个的定义和表达式见 2.3 节.

7.1.3 POTB 中的均匀性

当研究者感兴趣的主要对象为主效应, 即交互效应可忽略时, 主效应设计 (main effects plan, MEP) 在许多工程试验中就占有了重要的地位和应用. 一个正交的 MEP (orthogonal main effects plan) 可以保证所有主效应在估计时是互相不混杂的. 但是正交主效应设计的一个最大问题是其试验次数一般较大, 因此, 许多学者研究了替代方法, 比如, 近似正交设计 (Wang and Wu, 1992; Ma et al., 2000), 带有区组的主效应设计 (Mukerjee et al., 2002) 以及处理因子通过区组因子达到两两正交的主效应设计 (Bose and Bagchi, 2007; Bagchi, 2010) 等. 其中, Bagchi (2010) 得到了饱和的、$s^{3m}2^{3m}$ 水平、通过区组因子达到两两正交的主效应设计 (POTB). 最近, Chen 等 (2015b) 给出了一些直接地、可循环地构造混水平 POTB 设计的新方法.

表 7.1 中的设计 D_1 是一个含有 3 列, 水平为 3^32^3 的 POTB 设计. 通过水平置换 (因子 F_1 的水平 (0, 1, 2) 置换为 (0, 2, 1), 因子 F_3 的水平 (0, 1, 2) 置换为 (1, 0, 2)) 可以得到表 7.1 中的设计 D_2. 这里, 可以发现对于同水平的 POTB 设计, 其试验次数比正交主效应设计要小得多. 在 POTB 设计里, s 水平因子不与区组因子正交, 但通过区组因子可以使得处理因子之间两两正交.

考虑一个 N 行设计 D, 其包含一个区组因子 L 和多个处理因子 $F_1, F_2, \cdots, F_m$, 因子水平均为 s. 记一个 $N \times s$ 矩阵 $\boldsymbol{X}_{F_1}$ 为因子 F_1 的关联矩阵, 如果对于第 u 次试验, 因子 F_1 的水平为 i, 则该矩阵的第 (u, i) 个元素为 1, 否则为 0. 记矩阵 $\boldsymbol{M}_{F_1F_2}$ 为因子 F_1 对因子 F_2 的关联矩阵, 其第 (i, j) 个元素值表示因子 F_1 在 i 水平, 因子 F_2 在 j 水平的试验次数. 如果因子 F_1 和 F_2 满足下列条件:

$$\boldsymbol{M}_{F_1L}\boldsymbol{M}_{LF_2} = k\boldsymbol{M}_{F_1F_2}, \tag{7.4}$$

则称处理因子 F_1 和 F_2 可以通过区组因子达到正交, 其中 k 表示区组大小. 当设计 D 的任意一对因子都可以通过区组因子达到正交, 则称设计 D 是一个 POTB 设计.

表 7.1　两个 $3^3 2^3$ POTB 设计

区组因子	D_1			D_2		
	F_1	F_2	F_3	F_1	F_2	F_3
区组水平 1	0	0	0	0	0	1
	0	1	1	0	1	0
	1	0	1	2	0	0
	1	1	0	2	1	1
区组水平 2	1	1	1	2	1	0
	1	2	2	2	2	2
	2	1	2	1	1	2
	2	2	1	1	2	0
区组水平 3	2	2	2	1	2	2
	2	0	0	1	0	1
	0	2	0	0	2	1
	0	0	2	0	0	2

Chen 等 (2015a) 研究了水平置换和 POTB 设计之间的关系, 并得到了下面的定理.

定理 7.4　*如果 D_1 是一个 POTB 设计, 且设计 D_2 是由设计 D_1 中任意因子的水平置换后得到的, 那么设计 D_2 仍然是一个 POTB 设计.*

水平置换后, 设计的均匀性发生了变化, Chen 等 (2015a) 进一步给出了一个改进的优化算法来搜索因子水平置换下的最优或者近似最优的均匀 $s^{3m}2^{3m}$ POTB 设计. 下述算法步骤中 “CD” 为相应的均匀性度量值. 事实上, 搜索所有的 $(s!)^m$ 个水平置换后的设计 (算法步骤中用 $\mathcal{P}(D)$ 表示) 在 s 或者 m 增加时成为一个 NP- 难问题. 从文献 (Chen et al., 2015a) 中可见, 相应 R 程序代码在配置为 Intel Core(TM) i5-3230M Duo CPU 2.6GHz 4 GB 内存的一台个人电脑中运行, 最复杂的搜索情形也仅需几分钟可以完成 (循环次数 $\tau = 10000$).

算法　启发式搜索算法伪代码

1: 给定 τ (循环次数)
2: 给定 δ (门限) 和 $c_0 \in (0.1, 0.25)$
3: 采用 Chen 等 (2015c) 的算法 1, 产生一个初始的 $s^{3m}2^{3m}$ POTB 设计 D
　令 $D^{\min} = D^c = D$
4: **for** $i = 1$ 到 τ **do**
5: 　基于 D^c, 产生一个新的设计 $D^{\text{new}} \in \mathcal{P}(D)$
6: 　计算 $\nabla = \text{CD}^2(D^{\text{new}}) - \text{CD}^2(D^c)$, 并产生 c
　　$(c \sim U(0,1))$
7: 　**if** $(\nabla < 0)$ 或者 $(\nabla < \delta$ 且 $c < c_0)$ **then** 令 $D^c = D^{\text{new}}$
8: 　　**if** $\text{CD}^2(D^c) < \text{CD}^2(D^{\min})$ **then** 令 $D^{\min} = D^c$
9: **end for**

注 7.1 上述算法与 (Fang et al., 2006b) 中的算法存在两个不同点. 第一个是初始输入设计 D^c, 上述算法中的初始设计是由 POTB 设计开始的, 而非一个随机产生的 U-型设计. 第二个不同点是关于邻近设计 $\mathcal{P}(D)$ 的定义, 一个常规的 TA 搜索算法中是通过交换同一列的两个不同元素来得到邻近设计, 而 (Chen et al., 2015a) 的算法需要交换该列内两个不同水平的所有位置. 类似的应用可以参见 (Tang et al., 2012).

7.2 超饱和设计中的均匀性

超饱和设计 (supersaturated design, SSD) 是一类不足以估计所有由设计矩阵的列表示的主效应的部分因子试验设计. 在工业和科学试验中, 尤其在初始阶段, 经常需要考察大量的因子且试验次数因高额的费用而很有限. 然而, 在许多情形, 仅有少数因子被认为是有显著效应的. 基于*效应稀疏性原则*假设, SSD 能有效地识别主要的因子.

7.2.1 二水平 SSD 中的均匀性

已有大量的文献研究了二水平的 SSD. Booth 和 Cox (1962) 在 SSD 的首次系统构造中提出了$E(s^2)$ 准则, 此准则是在 m 个因子中只有两个活跃因子这一假设下对设计的非正交性的一个度量. 在 (Booth and Cox, 1962) 之后, 没有太多关于 SSD 的工作, 直到 (Lin, 1993) 发表 . 其他一些关于构造 $E(s^2)$-最优 SSD 的工作包括 (Nguyen, 1996; Cheng, 1997; Tang and Wu, 1997; Liu and Zhang, 2000; Butler et al., 2001; Liu and Dean, 2004; Bulutoglu and Cheng, 2004), 更多相关文献可以参见综述文章 (Georgiou, 2014).

令 $\boldsymbol{U} \in \mathcal{U}(n; 2^m)$, 当 $n < m+1$ 时, 此设计为二水平 SSD. 常用来比较 SSD 的 $E(s^2)$ 为

$$E(s^2) = \frac{2}{m(m-1)} \sum_{1 \leqslant i < j \leqslant m} s_{ij}^2,$$

其中 s_{ij} 是 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的 (i,j) 元素. Liu 和 Hickernell (2002b) 研究了 $E(s^2)$-最优和最小偏差的二水平 SSD 之间的联系. 当 $q = 2$ 时, 他们用 $a = 1+\beta$ 和 $b = 1+\beta\rho(\beta > 0, -1 \leqslant \rho < 1)$ 定义了 (2.50) 中的离散偏差, 并且证明了对于二水平因子设计, $E(s^2)$ 和离散偏差均可以表示为试验点间的 Hamming *距离* (或者*相遇数*) 的函数. 此关于 Hamming 距离的结论可以导出 $E(s^2)$ 和离散偏差的下界. 当一个设计 $\boldsymbol{U}$ 达到其中一个的下界, 则它也达到另一个的下界. 换句话说, 在离散偏差的意义下, $E(s^2)$-最优设计也是均匀的 (最小偏差). 他们进一步证明了哪些情况下此下界可以达到, 尽管此时离散偏差不等价于 $E(s^2)$ 准则.

定理 7.5 令 $\boldsymbol{U}$ 为一个具有 n 次试验、m 个因子的二水平设计, 其中每一列具有相同数量的 ± 1. 假设 $\rho\beta > -1$, 对 $e = -1, 0$ 或 1, $m = c(n-1)+e$. 同时, 假设

(a) n 是 4 的倍数且存在一个 $n \times n$ Hadamard 矩阵, 或者

(b) c 是偶数且存在一个 $2n \times 2n$ Hadamard 矩阵, 则 $E(s^2)$ 和离散偏差的下界均可达到.

相比 $E(s^2)$, 离散偏差更加一般化, 所以离散偏差是一个更加灵活的准则. 例如, $E(s^2)$ 忽略了两个或更多因子之间的交互效应. 但是, 离散偏差包含了所有可能阶的交互效应, 并且可以通过改变离散偏差定义中的 β 值来增强或减弱这些交互效应的重要性.

7.2.2 混水平 SSD 中的均匀性

二水平 SSD 可以用来筛选简单线性模型中的因子. 当因子和响应之间的关系是非线性或被多项式响应曲面模型近似时, 需要多水平或者混水平设计, 例如, 用来研究因子的非线性效应. 在过去二十多年里, 多水平或者混水平 SSD 的研究得到了极大的关注, 可参见 (Sun et al., 2011b; Liu Y and Liu M Q, 2011, 2012, 2013; Chen et al., 2013; Georgiou, 2014) 中的例子和有关参考文献. 我们知道所有 SSD 的列之间不可能完全正交. 许多作者考虑组正交性 (对设计对列进行分组, 其中每组中的列之间都是正交的). Fang 等 (2000a) 提出利用均匀设计和正交表来构造多水平 SSD 的方法.

文献中有许多比较多水平和混水平 SSD 的准则, 例如, aveχ^2 准则 (Yamada and Lin, 1999)、$\chi^2(D)$ 准则 (Yamada and Matsui, 2002)、$E(f_{\rm NOD})$ (Fang et al., 2003c) 准则、最小低阶矩混杂准则 (Xu, 2003) 和最小 χ^2 准则 (Liu et al., 2006). 特别地, 对一个设计 $\boldsymbol{U} \in \mathcal{U}(n; q_1 \times \cdots \times q_m)$, $E(f_{\rm NOD})$ 准则被定义为极小化

$$E(f_{\rm NOD}) = \sum_{1\leqslant i<j\leqslant m} f_{\rm NOD}^{ij} \Big/ \binom{m}{2}, \tag{7.5}$$

其中

$$f_{\rm NOD}^{ij} = \sum_{u=1}^{q_i}\sum_{v=1}^{q_j}\left(n_{uv}^{(ij)} - \frac{n}{q_i q_j}\right)^2,$$

$n_{uv}^{(ij)}$ 是第 i 和 j 列 (u,v) 对的个数, $n/(q_iq_j)$ 代表的是这两列中各水平组合的平均频数. 这里, 下标 NOD 代表设计的非正交性. 显然在正交设计中 $E(f_{\rm NOD}) = 0$. (Fang et al., 2003c) 证明了 $E(s^2)$ 和 ave χ^2 准则实际上是 $E(f_{\rm NOD})$ 准则的特殊情形且得到了 $E(f_{\rm NOD})$ 的一个下界, 此下界可作为设计最优性的一个基准. 同样地, 他们研究了 (2.52) 中的离散偏差和 $E(f_{\rm NOD})$ 的关系. (Fang et al., 2004a) 给出如下 $E(f_{\rm NOD})$ 的下界和达到此下界的充要条件, 其包含了 (Fang et al., 2003c) 中相应的下界和条件作为特例.

定理 7.6　令 $\boldsymbol{U} \in \mathcal{U}(n; q_1 \times \cdots \times q_m)$, 则

$$E(f_{\text{NOD}}) \geqslant \frac{n(n-1)}{m(m-1)}\left[(\gamma+1-\lambda)(\lambda-\gamma)+\lambda^2\right] + C(n, q_1, \cdots, q_m), \tag{7.6}$$

其中 $C(n, q_1, \cdots, q_m) = \dfrac{nm}{m-1} - \dfrac{1}{m(m-1)}\left(\sum_{i=1}^m \dfrac{n^2}{q_i} + \sum_{1 \leqslant i \neq j \leqslant m} \dfrac{n^2}{q_i q_j}\right)$, λ, γ 以及达到此下界的充要条件均与定理 2.23 中相同.

基于 (Fang et al., 2003c, 2004a) 这两篇文章的结论, 有如下结论.

定理 7.7　令 $\boldsymbol{U}$ 为一个 $\boldsymbol{U}$-型设计, 如果 $\boldsymbol{U}$ 中任何不同两列的 Hamming 距离的差异不超过 1, 则 $\boldsymbol{U}$ 既是离散偏差 (2.52) 下的一个均匀设计, 又是一个 $E(f_{\text{NOD}})$-最优设计.

该定理推广了 (Liu and Hickernell, 2002b) 中二水平因子设计的结论, 并且此结论导致对于任何一个 SSD, $E(f_{\text{NOD}})$ 最优性和 (2.52) 中的离散偏差有很强的联系. 作为特殊的等水平 SSD, $E(s^2)$-和 aveχ^2-最优 SSD 的均匀性可以通过上述定理直接得到.

Fang 等 (2003c, 2004a) 也提出了构造 $E(f_{\text{NOD}})$-最优设计和混水平均匀设计的方法. 更多 (2.52) 中离散偏差下的均匀 SSD 可以通过组合方法得到, 详细的方法见 3.6 节. 所有的结论均显示均匀性在度量和构造 SSD 中起到重要的作用.

7.3　分片拉丁超立方体设计中的均匀性

近些年, 同时含有定性、定量因子的计算机试验引起了越来越多的研究兴趣. 为了适应这类试验, Qian (2012) 首先提出了分片拉丁超立方体设计 (sliced Latin hypercube design, SLHD), 并给出了一个一般构造方法. 对于一个 $m \times q$ 的矩阵, 如果它的每一列都包含 m 个等距水平, 比如 $1, \cdots, m$, 则称之为拉丁超立方体设计, 记为 $L(m, q)$. 对于一个 $n \times q$ 的矩阵 S, 如果它能够分成 s 片, 且每一片在水平塌陷 $\lceil \cdot / s \rceil$ 后都是一个拉丁超立方体设计 $L(m, q)$, 则称之为有 s 片的 SLHD, 记为 SL(n, q, s), 这里 $m = n/s$, $\lceil a \rceil$ 是指大于等于 a 的最小整数. SLHD 遗传了拉丁超立方体设计的优良性质, 即整体设计本身以及每一片设计在任一维度上都具有投影均匀性, 达到最大分层. 关于 SLHD 的进一步研究, 主要包括具有列正交性或高于一维的投影均匀性的设计的构造, 例如, Yang 等 (2013)、Huang 等 (2014)、Cao 和 Liu (2015) 与 Yang 等 (2016) 构造了正交以及近似正交的 SLHD; Yin 等 (2014) 通过正交表构造了整体设计具有良好低维投影均匀性的 SLHD; Yang 等 (2014) 基于可分解正交表构造了整体设计与每一片设计都具有良好低维投影均匀性的 SLHD, 并且基于中心化 L_2 偏差准则提高了整体设计在试验区域中的均匀性.

SLHD 的一个缺点是它们在整个试验区域中不具有良好的均匀性, 这与计算机

试验需要的空间填充性是相违背的. 尽管 Yang 等 (2014) 考虑了该问题, 但是他们的设计严重依赖于可分解正交表的存在性, 而可分解正交表的试验次数和因子数都是受限的.

注意到在同时含有定性、定量因子计算机试验的应用中, SLHD 的每一片都对应于定性因子的一个水平组合. 当定性因子不同水平组合下的响应曲面相似时, 每个水平组合对应的设计点应该均匀地分布在试验区域中 (参见 (Huang et al., 2016). 整体设计的均匀性和每一片设计的均匀性没有一一对应的关系. 为了避免整体设计和每片设计的均匀性可能出现不一致的情况, 最近 Chen 等 (2016) 提出了结合两种均匀性的新优化准则, 基于此综合准则优化得到的设计称为均匀分片拉丁超立方体设计(uniform sliced Latin hypercube design, USLHD). 该设计不仅整体设计具有良好的均匀性, 而且每一层的设计点在试验区域中也均匀地分布.

7.3.1 一个综合的均匀性度量

Chen 等 (2016) 使用了中心化偏差 (CD) 衡量 SLHD 的均匀性. 对于一个 SLHD $D=(D_{(1)}^{\mathrm{T}},\cdots,D_{(s)}^{\mathrm{T}})^{\mathrm{T}}$, 用 $\mathrm{CD}(D)$ 和 $\mathrm{CD}(D_{(i)})$ 分别表示 D 和 $D_{(i)}$ 的 CD 值, 这里 $D_{(1)},\cdots,D_{(s)}$ 是 D 的 s 片. 一个综合的均匀性准则可以用以下式子表示

$$\mathrm{CD}(D,\xi)=\xi\mathrm{CD}(D)+(1-\xi)\sqrt[s]{\prod_{i=1}^{s}\mathrm{CD}(D_{(i)})}, \tag{7.7}$$

这里 $0\leqslant\xi\leqslant 1$ 是一个权重参数, $\sqrt[s]{\prod_{i=1}^{s}\mathrm{CD}(D_{(i)})}$ 是 $\mathrm{CD}(D_{(i)})(i=1,\cdots,s)$ 的几何平均值. 注意到 $\mathrm{CD}(D)$ 和 $\mathrm{CD}(D_{(i)})$ 可能具有不同的数量级, 这在比较两种均匀性以及决定合适的权重值 ξ 时可能会导致不公平. 因此, 不能像 (7.7) 式这样直接使用 CD 值, 而应使用一个整体设计和每片设计具有相同的数量级且能体现设计均匀性的一个度量. 为此, Chen 等 (2016) 引入了均匀性效率, 记设计 D 的均匀性效率为 $E_U(D)$, 其定义如下

$$E_U(D)=\frac{\mathrm{CD}(U)}{\mathrm{CD}(D)}, \tag{7.8}$$

这里, 设计 U 是基于 CD 下的均匀设计, 其试验次数和因子数与设计 D 相同. 很多均匀设计都可以在网站 `http://dst.uic.edu.hk/en/isci/uniform-design/uniform-design-tables` 上查到. 显然, $0<E_U(D)\leqslant 1$ 且 $E_U(D)$ 和 $E_U(D_{(i)})$ $(i=1,\cdots,s)$ 具有相同的数量级, 这里 $E_U(D)$ 和 $E_U(D_{(i)})$ 分别是 D 和 $D_{(i)}$ 的均匀性效率. 于是, Chen 等 (2016) 提出了如下形式的综合均匀性准则:

$$E_U(D,\omega)=\omega E_U(D)+(1-\omega)\sqrt[s]{\prod_{i=1}^{s}E_U(D_{(i)})}, \tag{7.9}$$

这里 $0 \leqslant \omega \leqslant 1$. 由于 $\mathrm{CD}(U)$ 是固定的, 因此 $E_U(D)$ 越大, $\mathrm{CD}(D)$ 就越小. 于是, 给定 ω, 目标就是在 $\mathcal{D}$ 中寻找设计 D^*, 使得

$$E_U(D^*, \omega) = \max_{D \in \mathcal{D}}(E_U(D, \omega)). \tag{7.10}$$

这里 D^* 称为一个 USLHD, $\mathcal{D}$ 是所有和 D 具有相同参数的 SLHD 的集合. 寻找 D^* 的优化算法将在 7.3.2 小节给出.

7.3.2 优化算法

为了得到 USLHD, 一个 SLHD 的“邻居” (neighbor) 是一个非常重要的概念, 将在优化算法中用到. 设 $\mathcal{D}$ 是包含所有 $\mathrm{SL}(n, q, s)$ 的集合, 则一个 $D_0 \in \mathcal{D}$ 的邻居可用下面的算法得到 (Chen et al., 2016).

算法 7.1

步骤 1 随机选择 D_0 的一列, 然后从该列的每一片中选择一个元素, 得到的 s 个元素记为 $a_1, \cdots, a_s$, 使得这 s 个元素满足 $\lceil a_1/s \rceil = \cdots = \lceil a_s/s \rceil$.

步骤 2 在 $a_1, \cdots, a_s$ 中随机选择两个元素, 记为 a_{i_1} 和 a_{i_2}, 交换它们的位置.

步骤 3 在 a_{i_1} 和 a_{i_2} 中随机选择一个, 比如 a_{i_1}, 再在与 a_{i_1} 同一列同一片元素中随机选择一个, 记为 b_1 $(b_1 \neq a_{i_1})$, 交换它们的位置.

得到的设计即为 D_0 的一个“邻居”. 显然, 算法中的交换程序并不改变一个 SLHD 的结构, 这是一个设计成为一个 SLHD“邻居”的必要条件.

为了搜索 USLHD, Chen 等 (2016) 使用了门限接受算法 (见 4.2 节). 下面给出了优化算法的详细步骤.

算法 7.2

步骤 1 给定参数 n, q, s, ω, 利用 Qian (2012) 提供的方法随机生成一个$\mathrm{SL}(n, q, s)$, 作为初始设计 D_0, 并计算 $E_U(D_0, \omega)$. 设置一个门限参数序列 $T_1 < \cdots < T_L = 0$, 在每个 $T_l (l = 1, \cdots, L)$ 下循环次数设置为 I, 设定两个循环指标 $l = 1,\ i = 1$.

步骤 2 利用算法 7.1, 随机构造 D_0 的一个“邻居”, 记为 D_{c}, 并计算 $E_U(D_{\mathrm{c}}, \omega)$.

步骤 3 如果 $E_U(D_{\mathrm{c}}, \omega) - E_U(D_0, \omega) \geqslant T_l$, 则用 D_{c} 替换 D_0, 否则保持 D_0 不变.

步骤 4 如果 $i \leqslant I$, 更新 $i = i + 1$, 转步骤 2.

步骤 5 如果 $l \leqslant L$, 更新 $l = l + 1$, 重置 $i = 1$, 转步骤 2, 否则输出 $D_{\mathrm{best}} = D_0$.

注 7.2 设置序列 $T_1, \cdots, T_L$ 是很重要的一步, 可以尝试几个序列, 然后选择一个收敛速度较快且能带来较大改进的序列.

利用算法 7.2 得到的设计称为权重为 ω 的综合均匀性准则下的 USLHD, 记为 $\mathrm{USL}(n, q, s, \omega)$. 它同时考虑了整体设计及其每一片设计的均匀性, 因此整体设计以及每一片的设计点在整个试验区域中都分布均匀.

依赖于初始设计的选择, 得到的设计 D_{best} 可能是局部最优的. 因此, 建议基于不同的初始设计, 运行算法多次, 最终在得到的设计中选择最好的一个. 还有一个很重要的问题, 就是对于一个 SL(n,q,s), 该给 ω 赋予什么样的具体值 ω_0, 才能使得得到的 USL(n,q,s,ω_0) 是所有 SL(n,q,s,ω) 中最有效的一个. 这将在 7.3.3 小节讨论.

7.3.3　权重 ω 的选择

由 (7.9) 和 (7.10) 式知, 当 ω 从 1 降到 0 时, $E_U(D^*)$ 随之减小, 而

$$\sqrt[s]{\prod_{i=1}^{s} E_U(D^*_{(i)})}$$

增大, 这里 $D^*_{(1)},\cdots,D^*_{(s)}$ 是 D^* 的 s 片. 权衡利弊, 更应该避免较低的 $E_U(D^*)$. 因此, Chen 等 (2016) 为 $E_U(D^*)$ 设置了一个下界 l_u, 然后最大化 $\sqrt[s]{\prod_{i=1}^{s} E_U(D^*_{(i)})}$. 这可引出如下的多限制优化问题

$$\max_{\omega} \sqrt[s]{\prod_{i=1}^{s} E_U(D^*_{(i)})}, \quad E_U(D^*) \geqslant l_u. \tag{7.11}$$

为了求解该问题, Chen 等 (2016) 引入了 ω- 迹, 即把

$$E_U(D^*) \quad 和 \quad \sqrt[s]{\prod_{i=1}^{s} E_U(D^*_{(i)})}$$

都看作是 ω 的函数, 作为纵轴, ω 作为横轴作图. 为了方便, Chen 等 (2016) 选取 $\{0,0.05,0.1,\cdots,1\}$ 作为 ω 的值. 当 ω 从 1 降到 0 时, $E_U(D^*)$ 将在某个 ω_0 处降到下界 l_u, 同时在所有 $E_U(D^*)$ 不小于 l_u 的 ω 中, $\sqrt[s]{\prod_{i=1}^{s} E_U(D^*_{(i)})}$ 将在 ω_0 处取到最大值. 也就是说, ω_0 就是我们要寻找的值. 下面是 (Chen et al., 2016) 中使用的例子.

例 7.1　假设 S_0 是一个随机生成的 SL$(18,3,3)$ 设计, 用算法 7.2 去搜索 USLHD, 其中 ω 从 0 到 1 以等距 0.05 共取 21 个值, 即 $\omega_i = 0.05(i-1), i=1,\cdots,21$, 门限参数序列 $T_1,\cdots,T_{11}$, 设置为 $T_i = -10^{-5}+10^{-6}(i-1), i=1,\cdots,11$. 对每一个 ω_i 都得到一个 USL$(18,3,3,\omega_i)$, 分别记为 $S_i, i=1,\cdots,21$, 然后计算 $E_U(S_i)$ 和 $\sqrt[3]{\prod_{j=1}^{3} E_U(S_{i(j)})}$, 这里 $S_{i(j)}(j=1,2,3)$ 是 S_i 的三片设计. 为了计算均匀性效率, 需要知道相应均匀设计 $U_{18}(18^3)$ 和 $U_6(6^3)$ 的 CD 值, 这两个值可以从前述网站里查到, 分别为 0.0506 和 0.1365. 进而, 可以计算 $E_U(S_i)$ 和 $\sqrt[3]{\prod_{j=1}^{3} E_U(S_{i(j)})}$ 的值, 其中 $i=1,\cdots,21$. 最后, 可以画出 ω-迹, 见图 7.1.

对于整体设计均匀性效率的下界 l_u, 随机生成 10000 个 SL$(18,3,3)$, 分别计算它们的均匀性效率, 然后选择上 5% 分位数作为 l_u 的值, 即 $l_u = 0.7314$. 在考虑整

体设计的 CD 值时, 该下界值可以保证得到的设计 USL$(18,3,3,\omega)$ 大约比 95% 随机生成的 SL$(18,3,3)$ 要小. 从图 7.1 可以看出, 当 ω 减小到 0.45 时, 整体设计的均匀性效率达到 l_u, 因此对于该例子, 可以设置 $\omega = 0.45$.

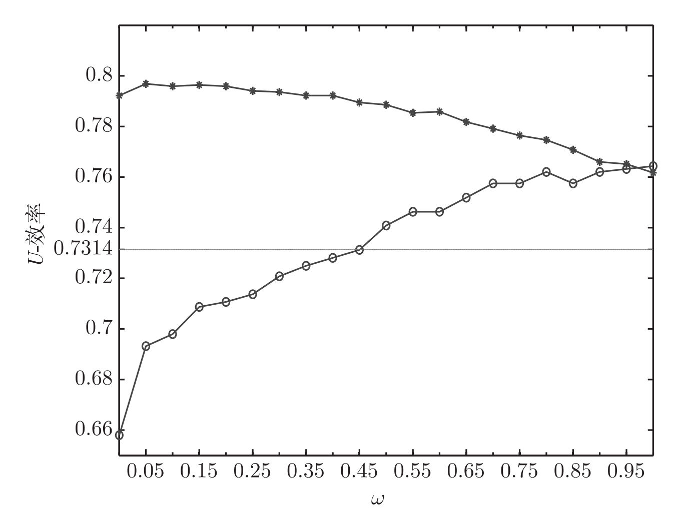


图 7.1 ω-迹图, 其中带有“$\circ$”的曲线对应的是整体设计的均匀性效率; 带有“$*$”的曲线对应的是三片设计几何平均值的均匀性效率

注 7.3 这里选取下界 l_u 值的方法不是本质的, 根据需要也可以选择其他的值, 比如上 2% 分位数, 这样可以保证在考虑整体设计的 CD 值时, 得到的设计 USL$(18,3,3,\omega)$ 大约比 98% 随机生成的 SL$(18,3,3)$ 要小.

注 7.4 根据图 7.1, 我们发现整体设计均匀性效率的下界 l_u 在 $\omega = 0.45$ 处取到, 而该值刚好在初始设置的 $\omega_1, \cdots, \omega_{21}$ 中, 即 $\omega_{10} = 0.45$. 一般地, 在使得整体设计均匀性效率大于下界 l_u 的所有 ω 值中, Chen 等 (2016) 建议取最小的那个 ω 值. 当没有足够的成本来计算 ω 的取值时, 为了能优化任何参数的 SLHD, Chen 等 (2016) 建议取 $\omega = 0.5$. 事实上 $\omega = 0.5$ 使得整体设计及其每片设计几乎有相等的重要性, 这看起来比较合理.

Chen 等 (2016) 进一步提供了两个同时含有定性、定量因子的计算机试验建立高斯过程模型的例子, 以检测 USLHD 的效果. 模拟结果显示, 当每片设计对应的真实响应曲面差别较大时, 整体设计更均匀的 SLHD 有更好的表现, 而当每片设计对应的真实响应曲面比较相似时, 每片设计更均匀的 SLHD 倾向于得到较小的均方预测误差平方根. 因此, 为了稳健起见, Chen 等 (2016) 建议 $\omega = 0.5$, 因为通常真实的响应函数是未知的.

7.4　水平值有误差时设计的均匀性

在真正做试验时, 实际实施的因子水平值可能带有误差, 见 (Box, 1963) 与 (Draper and Beggs, 1971). 考虑表 7.2 中的设计 $U(6;6^5)$ (来自于均匀设计网站 http://www.math.hkbu.edu.hk/UniformDesign/). 该设计的 WD^2 值为 0.2030. 当因子水平值带有误差时, 实际实施的因子水平变为 $u'_{ik}=u_{ik}+\varepsilon_{ik}$, 其中 ε_{ik} 为随机误差. 考虑误差服从具有有界支撑的均匀分布, 即 $\varepsilon_{ik}\overset{\text{i.i.d}}{\sim}\mathrm{Unif}(-\tau,\tau)$. 若取 $\tau=1/3$, 那么实际实施的设计可能如表 7.3 所示, 其 WD^2 值为 0.2182, 而原始 WD^2 值为 0.2030.

表 7.2　一个 $U(6;6^5)$ 设计

行	1	2	3	4	5
1	3	1	6	5	4
2	5	2	2	2	5
3	6	5	3	6	3
4	4	4	5	1	1
5	2	6	4	3	6
6	1	3	1	4	2

表 7.3　均匀误差 Unif$(-1/3,1/3)$ 下的 $U(6;6^5)$ 设计

行	1	2	3	4	5
1	2.6644	0.5613	5.8923	4.7964	3.4486
2	4.9332	1.5265	1.7120	1.7475	4.9865
3	5.8811	4.5608	2.5803	5.8632	2.5143
4	3.7633	3.6894	4.8018	0.9712	0.5016
5	1.8786	5.8181	3.6977	2.6817	5.9172
6	0.7735	2.5395	0.6299	3.9201	1.8249

Yang 等 (2010) 研究了因子水平值具有随机误差时设计的均匀性. 他们首先考虑了所有因子取值都有误差时设计的 WD 值, 然后研究了误差只出现在部分因子中的情况, 最后将该结果应用于均匀设计的构造上.

对任一 $U(n;n^s)$ 设计 $\boldsymbol{U}$, 当因子水平被均匀分布误差所污染时, 设计将变成 $\boldsymbol{Z}=\boldsymbol{U}+\boldsymbol{\varepsilon}$, 其中 $\boldsymbol{\varepsilon}=(\varepsilon_{ik})$, $\varepsilon_{ik}\overset{\text{i.i.d}}{\sim}\mathrm{Unif}(-\tau,\tau)$. 这种情况下, 产生的设计 Z 的 WD^2 值变为

$$\mathrm{WD}^2(\boldsymbol{Z})=-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\prod_{k=1}^s\left(\frac{3}{2}-|x_{ik}-x_{jk}+\delta_{ik}-\delta_{jk}|+|x_{ik}-x_{jk}+\delta_{ik}-\delta_{jk}|^2\right),$$

其中 $\delta_{ik}(=\varepsilon_{ik}/n)\overset{\text{i.i.d}}{\sim}\text{Unif}(-a,a)$, $a=\tau/n$. 令 $B_Z(n,s,a)=E(\text{WD}^2(\boldsymbol{Z})-\text{WD}^2(\boldsymbol{U}))$ 表示设计 $\boldsymbol{Z}$ 和设计 $\boldsymbol{U}$ 的 WD^2 值的期望偏差. Yang 等 (2010) 得出如下结论.

定理 7.8 对一个 $U(n;n^s)$ 设计 $\boldsymbol{U}$, 当 $a<1/(2n)$ 时, 有

(i) $\dfrac{n-1}{n}\left(\dfrac{5}{4}\right)^s\displaystyle\sum_{k=1}^{s}\binom{s}{k}\left(\dfrac{8}{15}a^2\right)^k<B_Z(n,s,a)<\dfrac{n-1}{n}\left(\dfrac{3}{2}\right)^s\displaystyle\sum_{k=1}^{s}\binom{s}{k}\left(\dfrac{4}{9}a^2\right)^k$;

(ii) 对任意固定的 n 和 s, $B_Z(n,s,a)$ 是 a 的增函数;

(iii) 对任意固定的 n 和 a, $B_Z(n,s,a)$ 是 s 的增函数.

定理 7.8 给出了设计的偏差在因子水平值有误差和没有误差时的不同. 结果显示, 有误差的设计平均来说不如没有误差的设计更均匀. 进一步, 误差越大, 平均意义来说设计的 WD 值就越大. 例如, 图 7.2 展示了表 7.2 中均匀设计的结果. 图 7.2(a) 中间的曲线显示了 $B_Z(n,s,a)$ 和 τ 值之间的关系. 这里模拟时, Yang 等 (2010) 用每个 τ 取 100 次重复的平均值作为 $B_Z(n,s,a)$. 上下两条曲线分别为 $B_Z(n,s,a)$ 的上下界. 为了解释 $B_Z(n,s,a)$ 的强度, Yang 等 (2010) 定义了该期望差别同 $\text{WD}^2(\boldsymbol{U})$ 值的相对比 (relative ratio, RR) 为 $\text{RR}=B_Z(n,s,a)/\text{WD}^2(\boldsymbol{U})$. 相对比越大意味着均匀性偏差改变越显著. 图 7.2 给出了 RR 值是如何随着 τ 的增加而变化的. 结果显示, 在误差相对较小 (如 $\tau<0.2$) 时均匀性偏差比较稳健. 这里的稳健指的是原始设计 (没有误差) 与受污染设计 (带有误差) 偏差改变的期望值不显著.

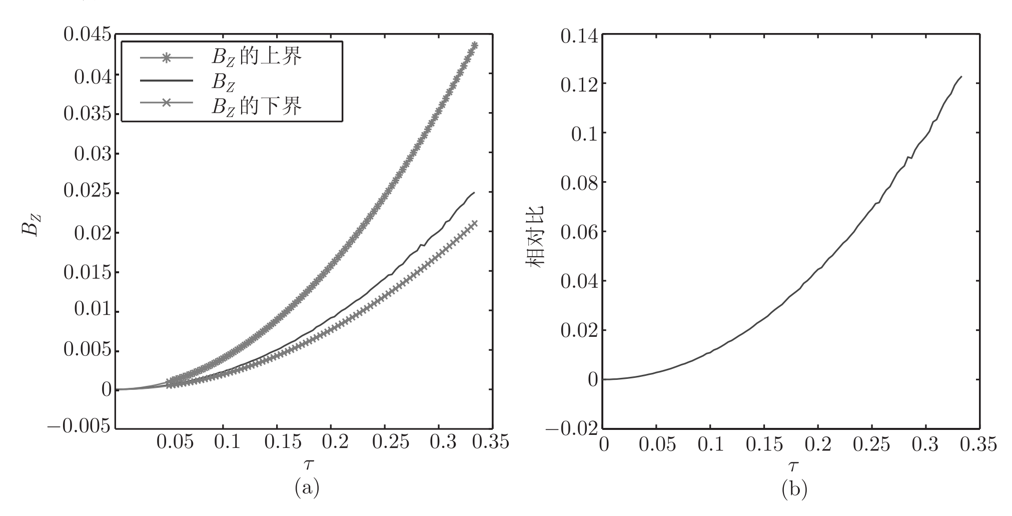


图 7.2 针对表 7.2 中设计的 (a) B_Z 对 τ 图; (b) RR 对 τ 图

Yang 等 (2010) 将该结果进一步推广, 考虑误差只出现在某些但不是所有因子中的情况. 不失一般性, 假设前 s_1 个因子在水平设置上存在误差, 并假设真实设计阵为 $\boldsymbol{Y}=\boldsymbol{U}+\boldsymbol{\varepsilon}$, 其中 $\boldsymbol{\varepsilon}=(\varepsilon_{ik})$ 且当 $k\leqslant s_1$ 时 $\varepsilon_{ik}\overset{\text{i.i.d}}{\sim}\text{Unif}(-\tau,\tau)$; 当 $k>s_1$ 时

$\varepsilon_{ik}=0$. 那么, 设计 $\boldsymbol{Y}$ 的 WD^2 值变成

$$\begin{aligned}\mathrm{WD}^2(\boldsymbol{Y})=&-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left[\prod_{k=s_1+1}^s\left(\frac{3}{2}-|x_{ik}-x_{jk}|+|x_{ik}-x_{jk}|^2\right)\right]\\&\times\left[\prod_{k=1}^{s_1}\left(\frac{3}{2}-|x_{ik}-x_{jk}+\delta_{ik}-\delta_{jk}|+|x_{ik}-x_{jk}+\delta_{ik}-\delta_{jk}|^2\right)\right],\end{aligned}$$

其中当 $k=1,\cdots,s_1$ 时 $\delta_{ik}=\varepsilon_{ik}/q_k$. 同时, 他们得到如下定理.

定理 7.9　对一个 $U(n;n^s)$ 设计 $\boldsymbol{U}$, 令 $B_Y(n,s,a,s_1)=E(\mathrm{WD}^2(\boldsymbol{Y})-\mathrm{WD}^2(\boldsymbol{U}))$, 则

(i) $$\frac{n-1}{n}\left(\frac{5}{4}\right)^s\sum_{k=1}^{s_1}\binom{s_1}{k}\left(\frac{8}{15}a^2\right)^k<B_Y(n,s,a,s_1)<\frac{n-1}{n}\left(\frac{3}{2}\right)^s\sum_{k=1}^{s_1}\binom{s_1}{k}\left(\frac{4}{9}a^2\right)^k;$$

(ii) 对任意固定的 n, s 和 s_1, $B_Y(n,s,a,s_1)$ 随着 a 的增大而增大;

(iii) 对任意固定的 n, s 和 a, $B_Y(n,s,a,s_1)$ 随着具有误差的因子个数 s_1 的增大而增大.

该定理表明, 带有误差的因子数越少且/或者误差越小, 设计的 WD 均匀性越好. 定理 7.8 和定理 7.9 的结果意味着, 如果误差比较小, 则传统的均匀设计相当稳健. 而且, 这些结果在均匀设计的构造上也非常有用. 在 $[0,1]^s$ 区域内构造均匀设计有两种框架. 传统的构造称作格子点抽样 (Patterson, 1954), 是选每个格子的中心作为试验点, 即 $x_{ik}=(u_{ik}-1/2)/n$, 其中 $1\leqslant i\leqslant n$, $1\leqslant k\leqslant s$. 另一种称作拉丁超立方体抽样 (McKay et al., 1979), 在格子内随机选点作为试验点, 即 $\tilde{x}_{ik}=(u_{ik}-\varepsilon_{ik})/n$, 其中 $\varepsilon_{ik}\overset{\text{i.i.d.}}{\sim}\mathrm{Unif}(0,1)$, $1\leqslant i\leqslant n$, $1\leqslant k\leqslant s$. Owen (1992b) 将这两种框架推广考虑正交表的空间填充性. Fang 等 (2002a) 在一维 ($s=1$) 情况下已证明, 格子点抽样最小化 $[0,1]$ 上的 WD 偏差. 定理 7.8 (ii) 对其进行推广, 在高维 ($s\geqslant 2$) 时表明, 从平均意义来讲, 该结果也是成立的.

需要注意的是, 以上结果是基于 WD 作为均匀性偏差讨论的. 实际上, 针对其他偏差如中心化 L_2 偏差, 类似的研究也可同样进行. 所得结果与定理 7.8 和定理 7.9 类似.

虽然只考虑了均匀随机误差, 但可以证明定理 7.8 和定理 7.9 的主要结果对其他随机误差结构同样成立. 比如对正态分布, 可以从理论上证明; 而对贝塔分布, 可以通过随机模拟来验证. 对正态分布而言, 因为其支撑集是无界的, 采用截断的正态随机误差可能更合适, 其中截断点的选择应该保证所有设计点都落在单位立方体内.

习　题

7.1 考虑例 3.11 中的可分解不完全区组设计.

1) 请写出它的关联矩阵 $\boldsymbol{Z}=(z_i^j)_{6\times 11}$.

2) 在 $\boldsymbol{Z}$ 中, 对于 z_i^j, r_i 和 t_j, $i=1,\cdots,6, j=1,\cdots,11$, 你能发现它们有什么关系?

3) 令 n_{ij}^{lh} 为 $\boldsymbol{Z}$ 的 $\boldsymbol{z}_i$ 和 $\boldsymbol{z}_j$ 两行中 (l,h)- 对的个数, 你能发现 n_{ij}^{lh} 有什么规律?

4) 请将你在 2) 和 3) 中的发现推广到更一般的 $\mathrm{RIB}(n,s,r,\boldsymbol{t},\boldsymbol{Z})$ 中.

7.2 请证明 (7.1) 中的表达式同样适用于 $\mathrm{RIB}(n,s,r,\boldsymbol{t},\boldsymbol{Z})$. 进而基于此表达式, 证明 (7.3) 是正确的.

7.3 令 δ_{ij} 为设计 $\boldsymbol{U}\in\mathcal{U}(n;q_1\times\cdots\times q_m)$ 中第 i 和 j 行的相遇数, 请用 δ_{ij} 表示 (7.5) 中的 $E(f_{\mathrm{NOD}})$.

7.4 假设 A, B 和 C 是如下三个 SL(12, 2, 2) 设计:

$$A=\left(\begin{array}{cccccc|cccccc} 11 & 8 & 1 & 6 & 4 & 10 & 7 & 12 & 3 & 2 & 9 & 5\\ 1 & 3 & 11 & 7 & 9 & 6 & 4 & 2 & 8 & 10 & 12 & 5 \end{array}\right)^{\mathrm{T}},$$

$$B=\left(\begin{array}{cccccc|cccccc} 9 & 1 & 5 & 8 & 12 & 4 & 2 & 3 & 10 & 11 & 7 & 6\\ 10 & 4 & 6 & 8 & 11 & 2 & 9 & 12 & 5 & 1 & 7 & 3 \end{array}\right)^{\mathrm{T}}$$

和

$$C=\left(\begin{array}{cccccc|cccccc} 1 & 7 & 12 & 4 & 9 & 6 & 5 & 10 & 2 & 8 & 3 & 11\\ 10 & 12 & 7 & 1 & 3 & 6 & 5 & 4 & 9 & 11 & 2 & 8 \end{array}\right)^{\mathrm{T}}.$$

1) 请计算 A, B 和 C 的 CD 值.

2) 令 $\xi=0.5$, 请计算 (7.7) 中 A, B 和 C 的综合 CD 值.

3) 请绘制 A, B 和 C 的散点图.

4) 通过 1)—3) 在设计的整体均匀性、综合均匀性以及整体 SLHD 和它的分片的均匀性关系中你有何发现?

7.5 考虑表 7.2 和表 7.3 中的两个设计.

1) 就 CD 值而言, 哪一个设计更好一些?

2) 令 $\tau=0.1,0.2,0.3,0.4$, 对每一个 τ, 在误差 $\mathrm{Unif}(-\tau,\tau)$ 下产生 50 个 $U(6;6^5)$ 设计, 计算这 50 个设计的 CD 值的平均值. 基于不同 τ 下的平均值和表 7.2 中 $U(6;6^5)$ 的 CD 值, 你有何发现?

第 8 章　混料均匀试验设计

本章将介绍混料试验和有约束的混料试验情形下的均匀设计及其建模方法. 首先将介绍一些经典的混料试验设计方法, 如 Scheffé **单纯形格子点设计** (scheffé simplex-lattice)、**单纯形中心设计**(simplex-centroid designs) 和**轴设计** 等. 然后介绍混料试验的均匀性测度和相应的均匀设计. 最后给出混料试验的一些建模方法.

8.1　混料试验简介

许多产品是由不同成分混合而成的. 例如, 建筑混凝土是最常用的建筑材料, 其由沙子、水和一种或多种水泥等成分混合而成. 负责混合这些成分的生产商或试验人员可能关心其硬度或强度等指标. 这些指标和水泥、沙和水的混合比例密切相关. 称决定混合比例的设计方案为**混料试验设计**, 其在很多领域中起到非常关键的作用, 例如, 化学工程、材料工程、制药工程等问题. 下面, 重新考虑例 1.6 如下.

例 8.1　*生产咖啡面包时, 需要考虑如下 11 种成分: 面粉、水、糖、植物起酥油、椰子薄片、盐、酵母、乳化剂、丙酸钙、咖啡粉和液体香精. 对于每个成分, 我们选择一个合适的比例使得最终的咖啡面包有好的味道. 因此, 需要一个如 (8.1) 式定义的单纯形 T^{11} 上的一个混料试验.*

设因子个数为 s, 第 i 个成分为 $x_i, i=1,\cdots,s$. 则试验区域为一个**单纯形**

$$T^s=\{(x_1,\cdots,x_s):x_j\geqslant 0,\ j=1,\cdots,s,\ x_1+\cdots+x_s=1\}. \tag{8.1}$$

一个混合 s 个成分, n 个试验点的设计是区域 T^s 上的 n 个点的集合. 根据约束 $x_1+\cdots+x_s=1$, 混料试验设计与无约束的因子试验设计是大不相同的.

然而, 在许多混料试验中, 这些成分还有更多的约束. 例如, 在例 8.1 中, 水和面粉是其主要成分, 而盐、糖和其他成分所在比例应该是比较低的. 因此, 更一般的约束是 $0\leqslant a_i\leqslant x_i\leqslant b_i\leqslant 1, i=1,\cdots,s$ 或 $\mathbf{0}\leqslant \boldsymbol{a}\leqslant \boldsymbol{x}\leqslant \boldsymbol{b}\leqslant \mathbf{1}$, 其中 $\boldsymbol{a}=(a_1,\cdots,a_s)$, $\boldsymbol{x}=(x_1,\cdots,x_s)$, $\boldsymbol{b}=(b_1,\cdots,b_s)$ 且 $\mathbf{0}$ 和 $\mathbf{1}$ 分别是元素都是 0 或 1 的向量. 因此, 试验区域变为

$$T^s(\boldsymbol{a},\boldsymbol{b})=\left\{\boldsymbol{x}:\mathbf{0}\leqslant \boldsymbol{a}\leqslant \boldsymbol{x}\leqslant \boldsymbol{b}\leqslant \mathbf{1},\sum_{i=1}^{s}x_i=1\right\}, \tag{8.2}$$

其为整个单纯形区域 T^s 中的子区域. 更一般的多成分混料试验有下面的约束

$$d_k \leqslant \sum_{i=1}^{s} a_{ki} x_i \leqslant e_k, \quad k = 1, \cdots, m. \tag{8.3}$$

在 $T^s(\boldsymbol{a}, \boldsymbol{b})$ 上的设计被称为有约束的混料试验设计, 将在 8.2.3 小节中讨论.

下面的例子说明, 我们需要在试验区域 $T^s(\boldsymbol{a}, \boldsymbol{b})$ 上考虑空间填充设计, 其已被 Piepel 等 (1993, 2002) 与 Borkowski 和 Piepel (2009) 所研究.

例 8.2 华盛顿里奇兰附近的汉福德建造的废物处理和固化计划 (Waste Treatment and Immobilization Plan, WTP), 其将核废料中的高放射性强度废物 (high-level waste, HLW) 和低活性废物 (low-activity waste, LAW) 以玻璃的形式固化后, 储存在地下大型储罐中. HLW 和 LAW 的玻璃必须满足废物负荷、化学耐久性、处理方式以及其他性能等要求. 为了满足这些要求, 我们通过试验设计方法, 测量感兴趣的玻璃性能, 并得到相应的模型以用于 WTP 操作. 科学家们希望通过试验设计达到如下目的: ① 鉴别出哪些能严重影响玻璃性能的成分; ② 确定感兴趣的试验区域 (表 8.1), 其需要确定在哪些区域是可行的, 哪些区域是不可行的. 因此, 我们根据试验区域上收集的试验数据, 通过建模来预测在哪些区域上的玻璃性能是可接受的.

表 8.1 例 8.2 中各成分以及相应的上下界约束

成分	名称	下界约束 (a_i)	上界约束 (b_i)
1	SiO_2	0.38	0.53
2	B_2O_3	0.05	0.14
3	Na_2O	0.04	0.15
4	Fe_2O_3	0.08	0.14
5	ZrO_2	0.00	0.06
6	MnO	0.00	0.05
7	SrO	0.00	0.05
8	Al_2O_3	0.04	0.085
9	Li_2O	0.02	0.06
10	CdO	0.0005	0.015
11	Spike	0.0015	0.015
12	NiO	0.001	0.01
13	Tl_2O_3	0.0002	0.002
14	Sb_2O_3	0.0002	0.002
15	SeO_2	0.0002	0.002
16	其他	0.042	0.042

在本章中, 分别讨论混料试验设计和有约束的混料试验设计. 最早的混料试验设计是由 Scheffé (1958) 基于二阶回归模型的单纯形格子点设计提出的. 后来,

Scheffé (1963) 提出单纯形中心设计. 我们把单纯形格子点设计和单纯形中心设计都称为 Scheffé 设计. 此后, 文献中提出各种各样的混料试验设计方法, Cornell (2002, 2011) 给出关于混料试验设计的综述. Chan (2000) 给出混料试验的最优设计的综述.

8.1.1 一些混料试验设计方法

在本小节中, 将介绍一些经典的混料试验设计方法, 如单纯形格子点设计、单纯形中心设计、轴设计和 Scheffé 型设计.

A. 单纯形格子点设计

为了估计在整个试验区域上的响应曲面, 一个自然的想法是选择一个设计使得其试验点能均匀地散布在整个单纯形区域.

考虑一个有 s 个成分的混料试验. 设 m 为一个正整数, 并假设每个成分等间隔地从 0 到 1 取 $(m+1)$ 个水平, 即

$$x_i = 0, 1/m, 2/m, \cdots, 1, \quad i = 1, \cdots, s. \tag{8.4}$$

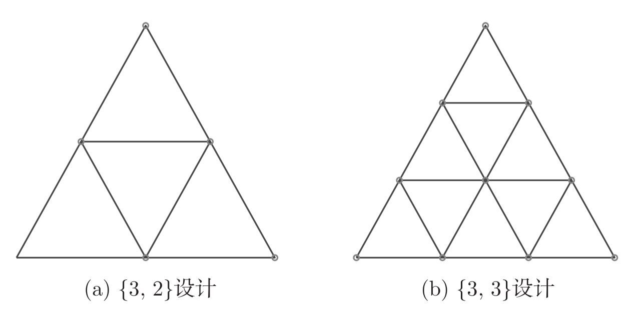
记一个 s 个成分的单纯形格子点设计为 $\{s, m\}$, 其有 $\binom{s+m-1}{m}$ 个设计点, 即其包括在约束 $x_1 + \cdots + x_s = 1$ 下所有可能的组合, 且每个成分的取值都来自于集合 $\{0, 1/m, 2/m, \cdots, 1\}$. 例如, 当 $s = 3$ 时, 单纯形格子点设计 $\{s, m\}$ 如下所示:

$m = 1$: 3 个设计点 (1, 0, 0), (0, 1, 0), (0, 0, 1);

$m = 2$: 6 个设计点 (1, 0, 0), (0, 1, 0), (0, 0, 1), (1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2);

$m = 3$: 10 个设计点 (1, 0, 0), (0, 1, 0), (0, 0, 1), (1/3, 2/3, 0), (1/3, 0, 2/3), (0, 1/3, 2/3), (2/3, 1/3, 0), (2/3, 0, 1/3), (0, 2/3, 1/3) (1/3, 1/3, 1/3).

这里的 $\{3, 2\}$ 和 $\{3, 3\}$ 单纯形格子点设计如图 8.1 所示. 从图 8.1 中可知, 许多点位于单纯形试验区域的边界上.



(a) {3, 2}设计　　(b) {3, 3}设计

图 8.1 $\{3, 2\}$ 和 $\{3, 3\}$ 单纯形格子点设计

B. 单纯形中心设计

一个 s 个成分的单纯形中心设计由下面的设计点构成:

(1) s 个由 $(1,0,0,\cdots,0)$ 的置换而得到的设计点,

(2) $\binom{s}{2}$ 个由 $(1/2,1/2,0,\cdots,0)$ 的置换而得到的设计点,

(3) $\binom{s}{3}$ 个由 $(1/3,1/3,1/3,0,\cdots,0)$ 的置换而得到的设计点,

$$\cdots\cdots$$

(4) 中心点 $(1/s,\cdots,1/s)$.

因此总的设计点数为 2^s-1. 例如, 当 $s=3$ 时, 设计点为 (1, 0, 0), (0, 1, 0), (0, 0, 1), (1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2) 和 (1/3, 1/3, 1/3). 与图 8.1(a) 中的单纯形格子点设计 $\{3,2\}$ 相比, 单纯形中心设计增加了一个中心点. 可以看到, 单纯形中心设计的大部分设计点都位于单纯形的顶点、边或面等边界上.

单纯形格子点设计和单纯形中心设计存在一些问题. 例如, 这些设计并没有为试验者提供比较灵活的设计方案, 且很多设计点都位于 T^s 的边界上. 然而, 在很多情形下, 边界上的设计点的某些元素为 0, 以至于无法进行试验. 因此, 这些边界上的设计点无意义. 为了克服这些缺点, 轴设计可能是一个合适的选择. 轴设计是一类设计点都在单纯形设计空间内部的混料设计. 另一种自然的想法是保持原设计的结构不变, 然后把设计点往 T^s 的中心点整体压缩, 例如, Fang 和 Wang (1994) 提出的 Scheffé 型设计.

C. 轴设计

我们称由单纯形区域 T^s 的各顶点与中心点 $(1/s,\cdots,1/s)$ 之间的线段为单纯形的轴. 中心点与每个顶点的距离为 $\sqrt{(s-1)/s}$. 设 d 为满足 $0<d<\sqrt{(s-1)/s}$ 的正数. 轴设计是在这 s 条轴上选择 s 个设计点的一种设计方法, 其要求中心点到这 s 个设计点的距离都是 d. 因此, 轴设计的每个设计点都位于单纯形的内部. 图 8.2 给出 $s=3$ 的一个轴设计. 显然, 不同的 d 为不同的轴设计. 最优的 d 值需要用其他准则确定. 将在 8.1.2 小节中讨论不同准则.

D. Scheffé 型设计

Fang 和 Wang (1994) 提出一种压缩方法来构造 Scheffé 型设计, 其核心思想是保持原设计的结构不变, 然后把边界点往 T_s 的中心压缩.

下面通过一个单纯形格子点设计 $\{3,3\}$ 的例子来说明这种压缩法的具体过程. 假设初始设计如图 8.1(b) 所示. 设 a 为某一待定的正数. 把这三个顶点做如下的

压缩:

$$
\begin{aligned}
(1,0,0) &\Rightarrow \left(1-\frac{1}{a}, \frac{1}{2a}, \frac{1}{2a}\right), \\
(0,1,0) &\Rightarrow \left(\frac{1}{2a}, 1-\frac{1}{a}, \frac{1}{2a}\right), \\
(0,0,1) &\Rightarrow \left(\frac{1}{2a}, \frac{1}{2a}, 1-\frac{1}{a}\right),
\end{aligned}
$$

其他边界点和内部的设计点可以类似地变化, 见图 8.3. 我们希望在某些准则下得到一个合适的 a.

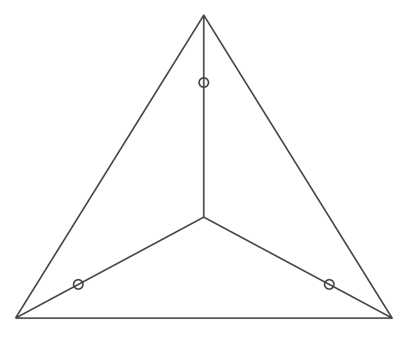


图 8.2　T^3 上的轴设计

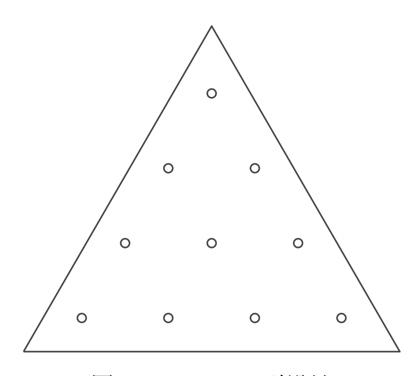


图 8.3　Scheffé 型设计

8.1.2　混料设计准则

在轴设计和 Scheffé 型设计中, 需要确定其最优参数. 本小节中, 我们给出一些基于距离和均匀性的准则.

Fang 和 Wang (1994) 提出 F-偏差来衡量混料设计的均匀性, 不过其不易计算. 因此, 他们建议用*均方距离* (mean square distance, MSD) 准则来衡量设计的优良性. Borkowski 和 Piepel (2009) 考虑*均方根距离* (root mean square distance, RMSD)、

平均距离 (average distance, AD) 以及**极大距离** (maximum distance, MD) 等准则. 令 $\mathcal{X}$ 为某一试验区域, 其可为 (8.1) 式定义的 T^s 或 (8.2) 式定义的 $T^s(\boldsymbol{a},\boldsymbol{b})$, 或其他区域. 设 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$ 为 $\mathcal{X}$ 上的一个设计. 随机向量 $\boldsymbol{x}$ 服从 $\mathcal{X}$ 上的均匀分布. 上述的准则可以定义如下:

均方距离和均方根距离

$$\mathrm{MSD}(\mathcal{P})=E[d^2(\boldsymbol{x},\mathcal{P})], \tag{8.5}$$

其中 $d(\boldsymbol{x},\mathcal{P})=\min_{1\leqslant j\leqslant n} d(\boldsymbol{x},\boldsymbol{x}_j)$, 且 $\mathrm{RMSD}(\mathcal{P})=\sqrt{\mathrm{MSD}(\mathcal{P})}$.

平均距离

$$\mathrm{AD}(\mathcal{P})=E[d(\boldsymbol{x},\mathcal{P})]. \tag{8.6}$$

极大距离

$$\mathrm{MD}(\mathcal{P})=\max_{\boldsymbol{x}} d(\boldsymbol{x},\mathcal{P}). \tag{8.7}$$

基于这些准则, 可以寻找一个最佳的设计使其最小化准则值. 当准则值不易计算时, 可以用蒙特卡罗方法寻找一个近似值. 设 $\boldsymbol{z}_1,\cdots,\boldsymbol{z}_N$ 为 $\mathcal{X}$ 上的 N 个随机样本, 则可用

$$\begin{gathered}
\mathrm{MSD}(\mathcal{P})\approx\frac{1}{N}\sum_{k=1}^{N} d^2(\boldsymbol{z}_k,\mathcal{P}),\\
\mathrm{AD}(\mathcal{P})\approx\frac{1}{N}\sum_{k=1}^{N} d(\boldsymbol{z}_k,\mathcal{P}),\\
\mathrm{MD}(\mathcal{P})\approx\max_{1\leqslant k\leqslant N} d(\boldsymbol{z}_k,\mathcal{P})
\end{gathered}$$

来近似这些准则值. 一般地, 在 T^s 上直接产生随机样本并不是一件容易的事, 有时可以借助间接方法来得到这些随机样本. Wang 和 Fang (1990) 提出**变换法**以生成 T^s 上的服从均匀分布的随机样本. 设 $\boldsymbol{y}=(y_1,\cdots,y_{s-1})$ 服从 $s-1$ 维单位立方体 $[0,1]^{s-1}$ 上的均匀分布. 令

$$\begin{cases}
z_i=(1-y_i^{\frac{1}{s-i}})\displaystyle\prod_{j=1}^{i-1} y_j^{\frac{1}{s-j}}, \quad i=1,\cdots,s-1,\\
z_s=\displaystyle\prod_{j=1}^{s-1} y_j^{\frac{1}{s-j}}.
\end{cases} \tag{8.8}$$

则 $\boldsymbol{z}=(z_1,\cdots,z_s)$ 服从 T^s 上的均匀分布.

Fang 和 Wang (1994) 对单纯形格子点设计 $\{3,3\}$ 和 $s=3$ 时的单纯形中心设计应用压缩法, 以得到 Scheffé 型设计. 基于 MSD 准则, 可以获得在压缩法中最优的参数 a, 其定义见 8.1.1 小节. 具体结果见表 8.2 从中可见压缩法可以减小 MSD 值. 此外, Prescott (2008) 给出另一种构造 Scheffé 型设计的方法, 其考虑对部分或全部成分加以上下界的约束, 即 $a_i \leqslant x_i \leqslant b_i, i=1,\cdots,s$. 这里忽略其具体细节, 有兴趣的读者可参考该文献.

表 8.2　Scheffé 设计和 Scheffé 型设计的 MSD

设计	Scheffé 设计的 MSD	Scheffé 型设计的 MSD	最优的 a
单纯形格子点设计 {3, 3}	0.03087	0.01568	4.836
单纯形中心设计	0.05553	0.02296	3.761

Scheffé 设计和 Scheffé 型设计共同的缺点是在于其设计点数大大受限. 例如, 对于 s 个成分的情形下, 单纯形格子点设计的试验点数为 $\binom{s+m-1}{m}$, 而单纯形中心设计的试验点数为 2^s-1. 在很多化学或工业试验中, 我们希望试验点数可以更灵活. 因此, Wang 和 Fang (1990) 提出**混料均匀设计** (uniform design for experiments with mixtures, UDEM) 或称之为**均匀混料设计** (uniform mixture design, UMD) (Borkowski and Piepel, 2009), 使得试验点可以均匀地散布在试验区域 T^s 中.

变换 (8.8) 可用于构造 T^s 上的均匀混料设计, 即首先构造一个 $[0,1]^{s-1}$ 上的均匀设计, 然后通过 (8.8) 获得 T^s 上的 n 个试验点的设计. 为了衡量这些混料设计的均匀性, 需要给出相应的偏差. 第 2 章给出的各种偏差不能直接应用到混料设计上. 8.2 节将讨论混料设计的均匀性度量以及相应的均匀混料设计的构造方法.

8.2　混料均匀设计

在本节中, 我们将介绍有约束和无约束这两类情形下的混料均匀设计.

8.2.1　混料设计的偏差

为了构造混料均匀设计, 需要首先给出相应的均匀性度量. 一般地, 有直接的和间接的这两类均匀性度量.

根据变换 (8.8), 衡量混料设计均匀性的一种间接方法是衡量在超立方体 C^{s-1} 上的设计的均匀性. 设 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$ 是单纯形 T^s 上的一个 n 个试验点的设计, 且 $\mathcal{P}_0=\{\boldsymbol{y}_1,\cdots,\boldsymbol{y}_n\}$ 为 C^{s-1} 上由 (8.8) 的逆变换而得到的设计. 则可以定义 $\mathcal{P}$ 的偏差为 $D(\mathcal{P}_0)$, 其中 $D(\cdot)$ 为第 2 章中的任一偏差类型. 因此, 可以选择一个 C^{s-1} 上的均匀设计 $\mathcal{P}_0$, 然后通过变换 (8.8) 获得单纯形上的均匀设计 $\mathcal{P}$. 不过这种间接的方法, 并不能非常准确地衡量混料设计的均匀性, 即当 $\mathcal{P}_0$ 是 C^{s-1} 上的

均匀设计时, 单纯形 T^s 上的相应设计 $\mathcal{P}$ 的均匀性可能并不太好.

Ning 等 (2011a) 给出另一种可以直接衡量混料设计均匀性的度量, DM_2-偏差, 这种偏差是直接定义在单纯形上的, 可以认为这种偏差是超立方体上的 L_2-星偏差的一种推广.

设 $\mathcal{P}=\{\boldsymbol{x}_1,\boldsymbol{x}_s,\cdots,\boldsymbol{x}_n\}$ 为单纯形 T^s 上一个 n 个试验点的设计. 记该设计 $\mathcal{P}$ 的设计矩阵为

$$\mathcal{P}=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1s} \\ x_{21} & x_{22} & \cdots & x_{2s} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{ns} \end{bmatrix}=(\boldsymbol{z}_1,\cdots,\boldsymbol{z}_s), \tag{8.9}$$

其中 $\boldsymbol{z}_i(i=1,2,\cdots,s)$ 是 $\mathcal{P}$ 的第 i 列. 定义该设计 $\mathcal{P}$ 的 DM_2-偏差如下:

$$\mathrm{DM}_2(\mathcal{P})=\left[\int_{T^s}\left|\frac{\mathrm{Vol}(R_M(\boldsymbol{x}))}{\mathrm{Vol}(T^s)}-\frac{N(\mathcal{P},R_M(\boldsymbol{x}))}{n}\right|^2\mathrm{d}\boldsymbol{x}\right]^{1/2}, \tag{8.10}$$

其中 $\mathrm{Vol}(A)$ 是区域 A 的体积, $N(\mathcal{P},R_M(\boldsymbol{x}))$ 是设计 $\mathcal{P}$ 落入区域 $R_M(\boldsymbol{x})$ 的点数, 其中区域 $R_M(\boldsymbol{x})(\boldsymbol{x}\in T^s)$ 的定义如下:

$$R_M(\boldsymbol{x})=\{\boldsymbol{t}=(t_1,\cdots,t_s):\boldsymbol{t}\in T^s \text{ 且 } t_i\leqslant x_i, i=2,3,\cdots,s\}. \tag{8.11}$$

当 $s=3$ 时, $R_M(\boldsymbol{x})$ 如图 8.4 所示.

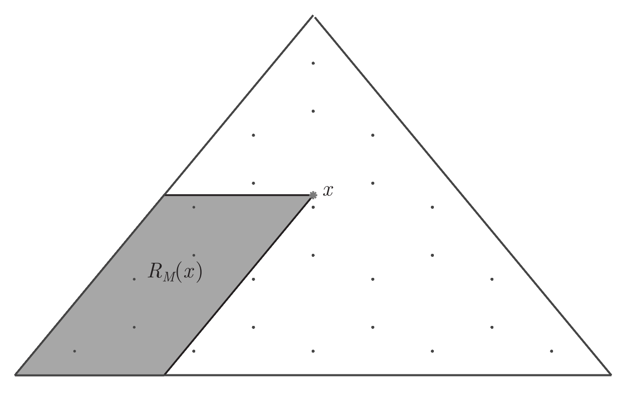


图 8.4 在 T^3 中由 x 确定的区域 $R_M(x)$

我们可以应用第 2 章中介绍的再生核希尔伯特空间的工具来获得 DM_2-偏差的计算表达式. 由核函数 (2.46) 的表达式, 可得

$$\mathcal{K}(\boldsymbol{z},\boldsymbol{t})=\int_{T^s}1_{R_M(\boldsymbol{x})}(\boldsymbol{z})1_{R_M(\boldsymbol{x})}(\boldsymbol{t})\mathrm{d}\boldsymbol{x}$$

$$=\frac{\sqrt{s}}{(s-1)!}\left[\max\left(1-\sum_{i=2}^{s}\max(z_i,t_i),0\right)\right]^{s-1}. \tag{8.12}$$

具体的证明细节可参见 (Ning et al., 2011a) 的附录 A.

设 $F_u(t)$ 为单纯形 T^s 上均匀分布的分布函数, 则 T^s 上均匀分布的密度函数为

$$f_u(t)=F_u'(t)=\begin{cases}1/\mathrm{Vol}(T^s), & t\in T^s,\\ 0, & \text{否则}.\end{cases}$$

把密度函数和 (8.12) 式代入 (2.45) 式, 可得 DM_2-偏差的计算表达式如下:

$$\begin{aligned}\mathrm{DM}_2(\mathcal{P})=&\left\{\int_{T^s\times T^s}\mathcal{K}(\boldsymbol{z},\boldsymbol{t})\mathrm{d}F_u(\boldsymbol{z})\mathrm{d}F_u(\boldsymbol{t})\right.\\ &\left.-\frac{2}{n}\sum_{i=1}^{n}\int_{T^s}\mathcal{K}(\boldsymbol{z}_i,\boldsymbol{t})\mathrm{d}F_u(\boldsymbol{t})+\frac{1}{n^2}\sum_{i=1,k=1}^{n}\mathcal{K}(\boldsymbol{z}_i,\boldsymbol{t}_k)\right\}^{1/2}\\ =&\left(\frac{\sqrt{s}}{(s-1)!}\right)^{1/2}\left\{C_{n,s}-\frac{2(s-1)!}{n}\sum_{i=1}^{n}\sum_{(\tau_2,\cdots,\tau_s)\in\{0,1\}^{s-1}}a_\tau\cdot(x_{i1})^{\tau_1}\cdot\prod_{j=2}^{s}x_{ij}^{\tau_j}\right.\\ &\left.+\frac{1}{n^2}\sum_{i=1,k=1}^{n}\left(\max\left(1-\sum_{j=2}^{s}\max(x_{ij},x_{kj}),0\right)\right)^{s-1}\right\}^{1/2},\end{aligned} \tag{8.13}$$

其中 $\{0,1\}^{s-1}=\{(t_1,\cdots,t_{s-1}):\ t_i=0 \text{ 或 } 1\}$, $C_{n,s}=\dfrac{((s-1)!)^3 2^{s-1}}{(2(s-1))!\prod_{k=0}^{s-2}(2s+k-1)}$, $a_\tau=\dfrac{(s-1)!}{(2(s-1)-\sum_{i=2}^{s}\tau_i)!}$ 且 $\tau_1=2(s-1)-\sum_{j=2}^{s}\tau_j$.

(8.13) 式的证明见 (Ning et al., 2011a) 的附录 B. 由公式 (8.13) 可知, 我们可以计算任意混料设计的偏差. 这个偏差的显式计算式有助于搜索混料均匀设计. 由公式 (8.13), 可知 DM_2-偏差具有下面的性质:

(1) 对于任意 (8.9) 中的混料设计矩阵 $\mathcal{P}$, 任意行置换都不改变其 DM_2 值.

(2) 对于任意混料设计矩阵 $\mathcal{P}$, 列置换基本不改变其 DM_2 值. 例如, 对于 $\{2,3,\cdots,s\}$ 的任意置换 $\{i_2,i_3,\cdots,i_s\}$, 记置换后的混料设计矩阵 $\mathcal{P}'=(\boldsymbol{z}_1,\boldsymbol{z}_{i_2},\cdots,\boldsymbol{z}_{i_s})$, 则 $\mathrm{DM}_2(\mathcal{P})=\mathrm{DM}_2(\mathcal{P}')$.

DM_2-偏差的上述性质说明对设计点的重排序并不改变其偏差值. 这种性质在实际中是很有用的, 因为可把试验顺序随机化.

8.2.2 混料均匀设计的构造方法

一个有 s 个成分的混料均匀设计是把这些点均匀地散布在区域 T^s 中. 基于 (8.8) 的变换方法, 构造均匀设计的步骤如下:

算法 8.1 (构造混料均匀设计的变换方法)

步骤 1 选择一个 $(s-1)$ 维的均匀设计 $U_n(n^{s-1})$, $\boldsymbol{U}=(u_{ki})$.

步骤 2 计算 $c_{ki}=(u_{ki}-0.5)/n$, 则

$$\boldsymbol{C}=\{\boldsymbol{c}_k=(c_{k1},\cdots,c_{k,s-1}),k=1,\cdots,n\}$$

是单位立方体 C^{s-1} 上的均匀分布.

步骤 3 计算

$$\begin{cases} x_{ki}=\left(1-c_{ki}^{\frac{1}{s-i}}\right)\prod_{j=1}^{i-1}c_{kj}^{\frac{1}{s-j}}, & j=1,\cdots,s-1,\\ x_{ks}=\prod_{j=1}^{s-1}c_{kj}^{\frac{1}{s-j}}, & k=1,\cdots,n. \end{cases} \tag{8.14}$$

则 $\mathcal{P}=\{\boldsymbol{x}_k=(x_{k1},\cdots,x_{ks}),k=1,\cdots,n\}$ 是单纯形 T^s 上的均匀设计.

在算法 8.1 的步骤 1 中, $s-1$ 维的均匀设计可由第 3 章或第 4 章的构造方法而得到, 如好格子点法或门限接受法等. 变换法是一种构造混料均匀设计的间接方法. 通常地, 这种方法能获得均匀性较好的设计, 而且这种方法易于操作.

例 8.3 构造一个 $n=11,s=3$ 的混料均匀设计. 表 8.3 的前两列为一个二维的均匀设计 $U_{11}(11^2)$, 记为 $\boldsymbol{U}_1$. 相应的设计 $\boldsymbol{C}=(c_{ki})$, $c_{ki}=(u_{ki}-0.5)/11$, 是一个 $[0,1]^2$ 上的均匀设计. 表达式 (8.14) 在 $s=3$ 的情形下有一个如下的简单形式:

$$\begin{cases} x_{k1}=1-\sqrt{c_{k1}},\\ x_{k2}=\sqrt{c_{k1}}(1-c_{k2}), & k=1,\cdots,n\\ x_{k3}=\sqrt{c_{k1}}c_{k2}, \end{cases} \tag{8.15}$$

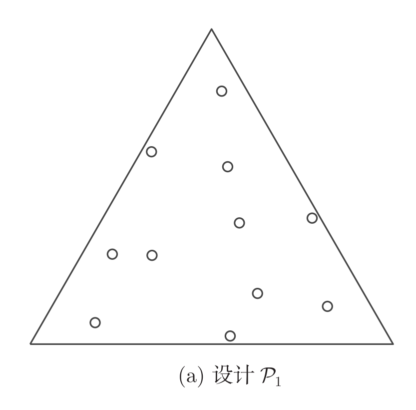
在 T^3 上相应的混料设计 $\mathcal{P}_1$ 为表 8.3 的第 5—7 列, 其图形见图 8.5(a). 该例子来自 (Fang and Ma, 2001a).

设 $\boldsymbol{C}_1$ 和 $\boldsymbol{C}_2$ 是 C^{s-1} 上的两个设计, $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 是由变换法得到的在 T^s 上相应的混料设计. 由于变换法是一个间接方法, 其可能出现一些不合理现象. 例如, 在 C^{s-1} 上设计 $\boldsymbol{C}_1$ 比 $\boldsymbol{C}_2$ 更均匀, 但在 T^s 上 $\mathcal{P}_2$ 比 $\mathcal{P}_1$ 更均匀.

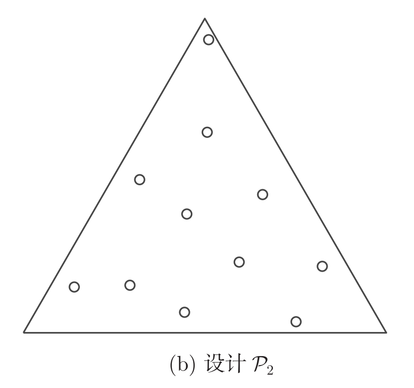
例 8.4 (例 8.3 续) 考虑表 8.3 中的两个设计 $\mathcal{P}_1$ 和 $\mathcal{P}_2$. 在 C^2 上相应的 $\boldsymbol{C}_1$ 和 $\boldsymbol{C}_2$ 也列在表 8.3 中. $\mathcal{P}_2$ 的设计点见图 8.5(b). 可计算 $\boldsymbol{C}_1$ 和 $\boldsymbol{C}_2$ 的混合偏差分别为 0.0692 和 0.0695. 不过 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 的 DM_2-偏差分别为 0.0486 和 0.0453, 即在 DM_2-偏差意义下, $\mathcal{P}_2$ 比 $\mathcal{P}_1$ 更均匀. 从图 8.5 上直观地看, $\mathcal{P}_2$ 更均匀一点.

表 8.3　$n = 11, s = 3$ 时的两个混料设计

$\boldsymbol{U}_1$		$\boldsymbol{C}_1$		$\mathcal{P}_1$			$\boldsymbol{U}_2$		$\boldsymbol{C}_2$		$\mathcal{P}_2$		
1	4	0.045	0.318	0.787	0.145	0.068	1	8	0.045	0.682	0.787	0.068	0.145
2	9	0.136	0.773	0.631	0.084	0.285	2	5	0.136	0.409	0.631	0.218	0.151
3	7	0.227	0.591	0.523	0.195	0.282	3	2	0.227	0.136	0.523	0.412	0.065
4	1	0.318	0.045	0.436	0.538	0.026	4	10	0.318	0.864	0.436	0.077	0.487
5	11	0.409	0.955	0.360	0.029	0.611	5	7	0.409	0.591	0.360	0.262	0.378
6	3	0.500	0.227	0.293	0.546	0.161	6	4	0.500	0.318	0.293	0.482	0.225
7	6	0.591	0.500	0.231	0.384	0.384	7	1	0.591	0.045	0.231	0.734	0.035
8	8	0.682	0.682	0.174	0.263	0.563	8	9	0.682	0.773	0.174	0.188	0.638
9	2	0.773	0.136	0.121	0.759	0.120	9	6	0.773	0.500	0.121	0.440	0.440
10	10	0.864	0.864	0.071	0.127	0.803	10	3	0.864	0.227	0.071	0.718	0.211
11	5	0.955	0.409	0.023	0.577	0.400	11	11	0.955	0.955	0.023	0.044	0.933



(a) 设计 $\mathcal{P}_1$



(b) 设计 $\mathcal{P}_2$

图 8.5　单纯形 T^3 上的两个混料设计

由例 8.4, 变换法可能不是一个最佳的构造混料均匀设计的方法. 可以考虑一些直接方法来构造混料均匀设计. 一般地, 也可以在 DM_2-偏差意义下, 直接应用一些随机优化算法来构造, 例如模拟退火算法和门限接受法等.

Ning 等 (2011b) 考虑一个构造混料均匀设计的直接方法. 该方法基于 NTLBG 算法 (Fang et al., 1994) 这一数值搜索法, 其可以应用于单纯形或其他试验区域上搜索均匀设计. NTLBG 算法联合数论方法和由 Linde 等 (1980) 提出的 LBG 算法, 后者可用于搜索椭球等高分布的代表点. 构造 n 个点的混料均匀设计的 NTLBG 算法的过程如下所示.

算法 8.2 (构造混料均匀设计的 NTLBG 算法)

步骤 1　给定试验次数 n 和因子个数 s, 选择一个正整数 $N \gg n$. 通过算法 8.1 的变换法生成一个 N 个点的近似均匀设计 $\mathcal{P}_N = \{\boldsymbol{t}_1, \cdots, \boldsymbol{t}_N\}$, 作为 T^s 上的训练样本.

步骤 2　在 T^s 上随机生成 n 个点作为初始设计 $\mathcal{P}_0 = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$. 计算其均匀性度量值 $D(\mathcal{P}_0)$.

步骤 3　把训练样本 $\mathcal{P}$ 剖分为 n 个子集, 即对于任意 $\boldsymbol{x}_i \in \mathcal{P}_0$, $i = 1, \cdots, n$,

$$P_{\boldsymbol{x}_i} = \left\{ \boldsymbol{t}_k : \ d(\boldsymbol{t}_k, \boldsymbol{x}_i) = \min_{\boldsymbol{x}_j \in \mathcal{P}_0} d(\boldsymbol{t}_k, \boldsymbol{x}_j) \right\}.$$

步骤 4　计算 $P_{\boldsymbol{x}_i}$ 的样本均值

$$\bar{\boldsymbol{x}}_i = \frac{1}{N_i} \sum_{\boldsymbol{t}_j \in P_{\boldsymbol{x}_i}} \boldsymbol{t}_j,$$

其中 N_i 是 $P_{\boldsymbol{x}_i}$ 中的元素个数. 令 $\mathcal{P}_{\rm new} = \{\bar{\boldsymbol{x}}_1, \cdots, \bar{\boldsymbol{x}}_s\}$ 为一个新设计并计算其均匀性度量值$D(\mathcal{P}_{\rm new})$.

步骤 5　对于事先给定的一个小的正数 α, 若 $D(\mathcal{P}_0) - D(\mathcal{P}_{\rm new}) > \alpha > 0$, 令 $\mathcal{P}_0 = \mathcal{P}_{\rm new}$ 并重复步骤 3 — 步骤 5. 否则中止算法并输出 $\mathcal{P}_{\rm new}$ 作为最终的设计.

算法 8.2 的一些解释如下. 在步骤 1 中, N 需远大于 n, 以保证算法的效果. 在步骤 2 中, 可以由算法 8.1 的方法生成一 n 个试验点的初始设计. 均匀性度量可以选取为 DM_2-偏差、均分距离、平均距离或极大距离. 在步骤 3 中, $\{P_{\boldsymbol{x}}, \boldsymbol{x} \in \mathcal{P}_0\}$ 有两个性质: ① $\bigcup_{\boldsymbol{x} \in \mathcal{P}_0} P_{\boldsymbol{x}} = T^s$; ② 对于任意 $i \neq j$, $P_{\boldsymbol{x}_i} \bigcap P_{\boldsymbol{x}_j} = \varnothing$. (Ning et al., 2011b) 说明 NTLBG 算法是构造混料均匀设计的有效方法.

8.2.3　有约束的混料均匀设计

在例 8.1 的生产咖啡、面包的混料试验中有 11 个成分, 不过水和面粉这两个成分所占的比例较大, 而糖、盐和其他成分所占的比例较小. 因此, 需要对成分的比例做更多的约束. 例如, $0 \leqslant a_i \leqslant x_i \leqslant b_1 \leqslant 1, i = 1 \cdots, s$ 或 $\boldsymbol{0} \leqslant \boldsymbol{a} \leqslant \boldsymbol{x} \leqslant \boldsymbol{b} \leqslant \boldsymbol{1}$, 其中 $\boldsymbol{a} = (a_1, \cdots, a_s)$, $\boldsymbol{b} = (b_1, \cdots, b_s)$, $\boldsymbol{0}$ 和 $\boldsymbol{1}$ 的元素分别都为 0 和 1. 则试验区域变为 (8.2) 中的 $T^s(\boldsymbol{a}, \boldsymbol{b})$. 易证, 区域 $T^s(\boldsymbol{a}, \boldsymbol{b})$ 非空的充要条件是

$$a \equiv \sum_{i=1}^{n} a_i < 1 < \sum_{i=1}^{n} b_i \equiv b. \tag{8.16}$$

上述条件可能包含一些过多的约束, 其可以精简为如下的约束:

$$a_i := \max(a_i, b_i + 1 - b), \quad b_i := \min(b_i, a_i + 1 - a). \tag{8.17}$$

因此, 对于第 i 个成分, 其下界和上界分别为 (8.17) 式中的 a_i 和 b_i.

在本小节中, 考虑在试验区域 $T^s(\boldsymbol{a}, \boldsymbol{b})$ 上构造均匀设计. Wang 和 Fang (1996) 应用变换法构造 $T^s(\boldsymbol{a}, \boldsymbol{b})$ 上的均匀设计. 但是, 当 $d_i = b_i - a_i$ 很小时, 这种方法效果不好. 为了克服变换法的这种缺点, Fang 和 Yang (2000) 采用条件法来构造有约束的

混料均匀设计. 关于条件法的更多讨论可参见 (Johnson, 1988). 设 $\boldsymbol{x}=(X_1,\cdots,X_s)$ 服从 $T^s(\boldsymbol{a},\boldsymbol{b})$ 上的均匀分布. 条件法基于以下事实:

(a) X_i 的边缘分布可以由一个简单的形式表出;

(b) 给定 $X_s=x_s^*$ 的条件后, $X_1,\cdots,X_{s-1}$ 的条件分布也是 $T^s(\boldsymbol{a}^*,\boldsymbol{b}^*)$ 上的均匀分布, 其中 $\boldsymbol{a}^*$ 和 $\boldsymbol{b}^*$ 通过某些计算式而得到.

下面给出上述性质 (a) 和 (b) 的一些具体公式. 设 $F(x_1,\cdots,x_s)$ 为 $(X_1,\cdots,X_s)$ 的多元分布函数. 蒙特卡罗方法中生成随机样本的条件法是基于下面的式子:

$$F(x_1,\cdots,x_s)=F_s(x_s)F_{s-1}(x_{s-1}|x_s)\cdots F_1(x_1|x_2,\cdots,x_s),$$

其中 $F_s(x_s)$ 是 X_s 的边缘分布, $F_{s-1}(x_{s-1}|x_s)$ 是给定 $X_s=x_s$ 时 X_{s-1} 的条件分布; $F_{s-2}(x_{s-2}|x_{s-1},x_s)$ 是给定 $X_{s-1}=x_{s-1},X_s=x_s$ 时, X_{s-2} 的条件分布, 以此类推. 产生服从 $F(x_1,\cdots,x_s)$ 的随机样本的步骤如下:

步骤 1　由分布 $F_s(x_s)$ 产生一个随机样本 x_s;

步骤 2　从条件分布 $F_{s-1}(x_{s-1}|x_s)$ 产生一个随机样本 x_{s-1}, 其中 x_s 为步骤 1 中的样本;

步骤 3　从条件分布 $F_{s-k}(x_{s-k}|x_s,\cdots,x_{s-k+1})$ 产生一个随机样本, x_{s-k}, 其中 $x_s,\cdots,x_{s-k+1}$ 为前面已获得的样本, $k=2,\cdots,s-1$.

步骤 4　输出随机样本 $\boldsymbol{x}=(x_1,\cdots,x_s)$.

显然, 在条件法中需要:

(1) 知道 $F_1,F_2,\cdots,F_s$ 的显式表达式;

(2) 容易获得 $F_s(x_s)$, $F_{s-1}(x_{s-1}|x_s),\cdots$, 和 $F_1(x_1|x_2,\cdots,x_s)$ 的随机样本.

基于如下的结果, 可以应用条件法获得在单纯形 T^s 的均匀分布的随机样本:

(A) X_s 的边缘分布为

$$F_s(x)=\int_0^x(s-1)(1-y)^{s-2}\mathrm{d}y=1-(1-x)^{s-1}.$$

(B) 变换. 设 $u_2,\cdots,u_s$ 为 $s-1$ 个随机数, 即其为均匀分布 $U(0,1)$ 上的随机样本. 令

$$\begin{aligned}
&x_s=1-(1-u_s)^{\frac{1}{s-1}},\\
&x_{s-i}=\left\{1-(1-u_{s-i})^{\frac{1}{s-i-1}}\right\}\left(1-\sum_{j=0}^{i-1}x_{s-j}\right),\quad i=1,\cdots,s-2,\\
&x_1=1-\sum_{i=2}^{s}x_i.
\end{aligned}$$

则 $(x_1,\cdots,x_s)$ 是单纯形 T^s 上的均匀分布 $U(T^s)$ 的随机样本, 且 $(x_2,\cdots,x_s)$ 服从 $S^{n-1}=\{\boldsymbol{x}:\boldsymbol{x}\in R_+^s,\sum_{i=1}^s x_i\leqslant 1\}$ 上的均匀分布.

(C) 条件分布 $F(x_1,\cdots,x_{s-1}|x_s)$ 为下面区域上的均匀分布

$$\begin{aligned}T^{s-1}(1-x_s)=&\{(x_1,\cdots,x_{s-1}):0\leqslant x_j\leqslant 1-x_s,\\&j=1,\cdots,s-1,x_1+\cdots+x_{s-1}=1-x_s\}.\end{aligned}$$

由这些性质说明, 可以对 $F(x_1,\cdots,x_{s-1}|x_s)$ 迭代 (A), (B) 和 C), 从而获得条件分布 $F_{s-1}(x_{s-1}|x_s)$. 进一步地, 分布 $F(x_1,\cdots,x_{s-2}|x_s,x_{s-1})$ 也是

$$\begin{aligned}T^{s-2}(1-x_s-x_{s-1})=&\{(x_1,\cdots,x_{s-2}):0\leqslant x_j\leqslant 1-x_s-x_{s-1},\\&j=1,\cdots,s-2,x_1+\cdots+x_{s-2}=1-x_{s-1}-x_s\}\end{aligned}$$

上的均匀分布. 上述过程可以一直迭代, 直到获得 $F_1(x_1|x_2,\cdots,x_s)$.

若 $\boldsymbol{x}$ 服从 $T^s(\boldsymbol{a},\boldsymbol{b})$ 上的均匀分布, 上述方法都可以使用. 设 $\boldsymbol{b}^*=(\boldsymbol{b}-\boldsymbol{a})/(1-a)$, $\boldsymbol{y}=(Y_1,\cdots,Y_s)=(\boldsymbol{x}-\boldsymbol{a})/(1-a)$, 其中 a 如 (8.16) 式定义. 则 $\boldsymbol{y}$ 服从 $T^s(\boldsymbol{0},\boldsymbol{b}^*)$ 上的均匀分布. 不失一般性, 可设 $\boldsymbol{a}=\boldsymbol{0}$. 因此, 只需关注于 $T^s(\boldsymbol{0},\boldsymbol{b})$ 上服从均匀分布的随机样本. 可知, Y_s 的边缘分布为

$$F_{Y_s}(y)=\frac{P(d_s\leqslant X_s\leqslant y)}{P(d_s\leqslant X_s\leqslant b_s)}=\begin{cases}1, & y\geqslant b_s,\\ \dfrac{(1-d_s)^{s-1}-(1-y)^{s-1}}{(1-d_s)^{s-1}-(1-b_s)^{s-1}}, & d_s<y<b_s,\\ 0, & y\leqslant d_s,\end{cases}$$

其中 $d_s=\max(0,1+b_s-b)$. 设 u 服从 $U(0,1)$. 由蒙特卡罗方法中逆变换方法, 可知

$$y_s\equiv F_{Y_s}^{-1}(u)=1-[u(1-b_s)^{s-1}+(1-u)(1-d_s)^{s-1}]^{1/(s-1)}$$

是 $F_{Y_s}(y)$ 的一个随机样本. 给定 $Y_s=y_s$, 条件分布 $(Y_1,\cdots,Y_{s-1})$ 服从 $T^{s-1}(\boldsymbol{0},\boldsymbol{b}_{(1)})$ 上的均匀分布, 其中 $\boldsymbol{b}_{(1)}=(b_1,\cdots,b_{s-1})/(1-y_s)$. 因此, 可以应用上述方法来得到 $T^{s-1}(\boldsymbol{0},\boldsymbol{b}_{(1)})$ 上的均匀分布.

令

$$\begin{aligned}&\Delta_s=1,\quad \Delta_k=1-\sum_{i=k+1}^{s}y_i,\quad k=s-1,\cdots,2,\\&d_k=\max\left\{a_k/\Delta_k,1-\sum_{i=1}^{k-1}b_i/\Delta_i\right\},\quad k=s,s-1,\cdots,2,\\&\Phi_k=\max\left\{b_k/\Delta_k,1-\sum_{i=1}^{k-1}a_i/\Delta_i\right\},\quad k=s,s-1,\cdots,2,\end{aligned}$$

$$G(u,d,b,c,k)=c\{1-[u(1-b)^k+(1-u)(1-d)^k]^{1/k}\}.$$

若产生 $s-1$ 个服从 $U(0,1)$ 的随机数 $u_2,\cdots,u_s$, 并记

$$x_k=G(u_k,d_k,\Phi_k,\Delta_k,k-1),\quad k=s,s-1,\cdots,2,$$
$$x_1=1-\sum_{k=2}^{s}x_k,$$

则 $\boldsymbol{x}=(x_1,\cdots,x_s)$ 服从 $T^s(\boldsymbol{a},\boldsymbol{b})$ 上的均匀分布的随机样本. 具体证明参见 (Fang and Yang, 2000). 因此, 构造在 $T^s(\boldsymbol{a},\boldsymbol{b})$ 上有约束的混料均匀设计的过程如下所示.

算法 8.3 (构造有约束的混料均匀设计的条件法)

步骤 1　选择一个超立方体中的 $s-1$ 维的均匀设计 $U_n(n^{s-1})$, 并记其为 $\boldsymbol{U}=(u_{ij})$;

步骤 2　计算

$$t_{i,j+1}=\frac{u_{ij}-0.5}{n},\quad i=1,\cdots,n,\ j=1,\cdots,s-1;$$

步骤 3　对每个 i, 计算

$$x_{ik}=G(t_{ik},d_k,\Phi_k,\Delta_k,k-1),\quad k=s,s-1,\cdots,2,$$
$$x_{i1}=1-\sum_{k=2}^{s}x_{ik}.$$

则 $\{\boldsymbol{x}_i=(x_{i1},\cdots,x_{is}),i=1,\cdots,n\}$ 是 $T^s(\boldsymbol{a},\boldsymbol{b})$ 上的均匀设计.

例 8.5　该例子来自 (Fang, 2002). 在某一药物研究中, 我们希望将轻微的极性药物溶解在水, 以及乙醇和丙二醇这两种助溶剂的混合物中, 以提高药物的溶解度. 试验人员希望知道药物在溶剂混合物中的溶解度曲线是否存在和哪里存在最大值. 响应值用蒸汽压力 (y, mmHg) 来测量. 乙醇 (x_1)、丙二醇 (x_2) 和水 (x_3) 这三个因子的取值范围为

$$T^3(\boldsymbol{a},\boldsymbol{b})=\{\boldsymbol{x}=(x_1,x_2,x_3):0.0463\leqslant x_1\leqslant 0.7188,$$
$$0.0272\leqslant x_2\leqslant 0.5776,0.2272\leqslant x_3\leqslant 0.9265,x_1+x_2+x_3=1\}.$$

考虑在 $T^3(\boldsymbol{a},\boldsymbol{b})$ 上 $n=12$ 的有约束的混料均匀设计. 这 12 个设计点和相应的相应值分别如表 8.4 所示.

此外, Jing 等 (2007) 应用本节的方法来研究一个实际问题, 其考虑由变色栓菌的固态发酵而生产漆酶. 在他们的例子中, 选择了一个 $U_{15}(5^4)$. 更多的细节见相应的文献.

表 8.4 例 8.5 中的设计和响应

编号	x_1	x_2	x_3	y
1	0.12296	0.11323	0.76381	31.5042
2	0.35453	0.04379	0.60168	51.5640
3	0.24176	0.36565	0.39259	30.0888
4	0.26535	0.47616	0.25849	28.9297
5	0.05914	0.41827	0.52258	21.5290
6	0.55981	0.07786	0.36233	56.6437
7	0.52303	0.22904	0.24793	44.9973
8	0.20319	0.18860	0.60821	36.6152
9	0.11576	0.54129	0.34296	20.3480
10	0.40656	0.15008	0.44336	45.6282
11	0.14105	0.27173	0.58723	30.8623
12	0.33631	0.31707	0.34662	37.0407

8.2.4 不规则区域上的均匀设计

在有些混料试验中, 对于成分 $x_1, \cdots, x_s$ 有更多的约束, 以至于其试验区域变为 R^s 中的不规则区域.

例 8.6 考虑 (Borkowski and Piepel, 2009) 中的一个问题. 某个产品的生产过程中需要考虑三个成分 x_1, x_2 和 x_3, 它们有如下的约束:

$$0.1 \leqslant x_1 \leqslant 0.7, \quad 0 \leqslant x_2 \leqslant 0.8, \quad 0.1 \leqslant x_3 \leqslant 0.6 \quad 且 \quad x_1 - x_2 \geqslant 0.$$

相应的试验区域 S 如图 8.6 (a) 所示, 即其试验区域由点 C, D, E, F, G 连接而成, 其中等边三角形的三个顶点 $(1,0,0)$, $(0,1,0)$ 和 $(0,0,1)$ 构成 $T^3 = \{(x_1,x_2,x_3)|x_1 + x_2 + x_3 = 1, x_i \geqslant 0\}$. 对于二维平面上的区域 S, 可以用等距变换把它变换至 R^2 上, 具体变换公式将在后面给出. 图 8.6 (b) 给出 S 在等距变换后的图形, 并记为 S_1, 其由点 C_1, D_1, E_1, F_1, G_1 连接而成. 此时, 单纯形 T^3 的三个顶点分别变出二维中的 $(0,0)$, $(\sqrt{2},0)$ 和 $(\sqrt{2}/2, \sqrt{6}/2)$.

A. 均匀性度量

为了衡量在不规则区域 S 上设计点的均匀性, 8.2.3 小节中介绍的 DM_2-偏差已不适合. 此时, 可以考虑 (8.5)—(8.7) 式中定义的均分距离、平均距离和极大距离等. 然而, 这些距离并没有显式表达式. 为此, Chuang 和 Hung (2010) 提出**中心复合偏差** (central composite discrepancy, CCD), 用于测量 S 上设计点的均匀性.

对于 S 中的任意固定点 $\boldsymbol{x} = (x_1, \cdots, x_s)$, 则通过点 $\boldsymbol{x}$ 且垂直于第 i 条轴的 $(s-1)$ 维超平面可以把第 i 条轴分为两部分: $(-\infty, x_i]$ 和 $(x_i, +\infty)$, $i = 1, \cdots, s$. 通过这种方法, 可以把区域 S 剖分为 2^s 个小区域, 分别记为 $S_1(\boldsymbol{x}), \cdots, S_{2^s}(\boldsymbol{x})$, 其

中点 $\boldsymbol{x}$ 为这些子区域的中心. 设 $\mathcal{P}$ 为 S 上的 n 个点设计, 则 $\mathcal{P}$ 的 CCD 定义如下:

$$\mathrm{CCD}_p(n,\mathcal{P})=\left\{\frac{1}{V(S)}\int_S\frac{1}{2^s}\sum_{t=1}^{2^s}\left|\frac{N(S_t(\boldsymbol{x}),\mathcal{P})}{n}-\frac{V(S_t(\boldsymbol{x}))}{V(S)}\right|^p\mathrm{d}\boldsymbol{x}\right\}^{1/p},\tag{8.18}$$

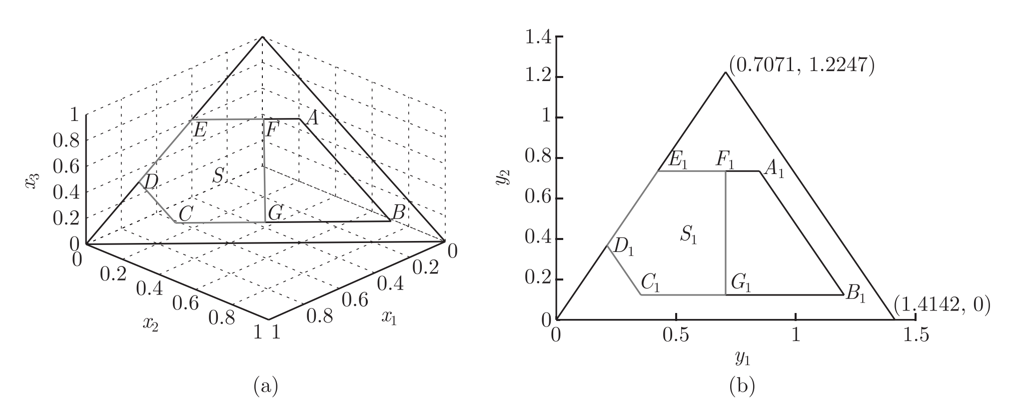


图 8.6　例 8.6 中相应图形: (a) 由点 C, D, E, F, G 连接而成的试验区域 S; (b) 等距变换后的 S 的图形 S_1, 其由点 C_1, D_1, E_1, F_1, G_1 连接而成

其中 $p>0$, $N(S_t(\boldsymbol{x}),\mathcal{P})$ 表示设计 $\mathcal{P}$ 落入 $S_t(\boldsymbol{x})$ 的点数, $V(S_t(\boldsymbol{x}))$ 和 $V(S)$ 分别表示 $S_t(\boldsymbol{x})$ 和 S 的体积. 显然, $\mathrm{CCD}_p(n,\mathcal{P})$ 值越小说明该设计的均匀性越好. 记 $Z(n)$ 为 S 上所有 n 个点的设计的集合. 给定试验次数 n, 定义 S 上的均匀设计 $\mathcal{P}^*$ 如下

$$\mathcal{P}^*=\arg\min_{\mathcal{P}\in Z(n)}\mathrm{CCD}_p(n,\mathcal{P}).$$

类似于超立方体中设计的 CD, WD 和 MD, (8.18) 式中的 p 值通常取为 2. 在实际应用中, 体积 $V(S_t(\boldsymbol{x}))$ 和 $V(S)$, 以及区域 S 上的积分往往不易计算. 为此, 可以用 N 个点代表连续区域 S, 其中 $N\gg n$, 则 (8.18) 式的 $\mathrm{CCD}_p(n,\mathcal{P})$ 值可用下面的表达式近似:

$$\mathrm{CCD}_2(n,\mathcal{P})\approx\left\{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{2^s}\sum_{t=1}^{2^s}\left|\frac{N(S_t(x_i),\mathcal{P})}{n}-\frac{N(S_t(x_i))}{N}\right|^2\right\}^{1/2}.\tag{8.19}$$

这 N 个点可取为 N 个点的近似均匀设计或 S 上的格子点.

B. **构造方法**

给定 CCD 等均匀性度量后, 在不规则区域 S 上构造均匀设计并不是一件很容易的事情. 为此, Chuang 和 Hung (2010) 采用交换算法, Chen 等 (2014) 采用离散粒子群优化算法等来构造不规则区域上的均匀设计. 此外, Liu Y 和 Liu M Q (2016) 提出一种构造复杂约束的混料均匀设计的构造算法.

Chuang 和 Hung (2010) 证明交换算法有较少的迭代次数和较快的收敛速度, 其比穷举法可以节约大量的运算时间, 且能找到均匀性很好的设计. 交换算法的具体过程如下所示.

算法 8.4 (交换算法)

步骤 1 选取一个 S 上的 N 个点的近似均匀设计 $\mathcal{P}=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N\}$, 其中 $N\gg n$.

步骤 2 任意选择试验区域 $\mathcal{P}$ 中的 n 个点作为初始的当前设计 "Cdesign", 例如选择 Cdesign $=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$; 令迭代计数器 $i=0$, 且 Ndesign = Cdesign.

步骤 3 While $i=0$ 或 Ndesign $\neq$ Cdesign

令 $i=i+1$, Cdesign=Ndesign;

for $j=1$ to n do

令 $\boldsymbol{x}^*=\arg\min_{\boldsymbol{x}\in\mathcal{P}\backslash\mathrm{Ndesign}}\mathrm{CCD}_p(n,\{\boldsymbol{x}\}\bigcup\mathrm{Ndesign}\backslash\{\boldsymbol{x}_j\})$;

if $\mathrm{CCD}_p(n,\{\boldsymbol{x}^*\}\bigcup\mathrm{Ndesign}\backslash\{\boldsymbol{x}_j\})\leqslant\mathrm{CCD}_p(n,\mathrm{Ndesign})$

令 Ndesign $=\{\boldsymbol{x}^*\}\bigcup\mathrm{Ndesign}\backslash\{\boldsymbol{x}_j\}$;

end if

end for

end while.

步骤 4 输出 Cdesign, $\mathrm{CCD}_p(n,\mathrm{Cdesign})$ 和 i.

在算法 8.4 的步骤 1 中, 更大的 N 可以获得更优的输出设计, 但需要更多的迭代时间. 一般地, 我们选一个 N 使得 $N/n>5$. 在算法 8.4 的步骤 3 中, 记号 $\mathrm{Ndesign}\backslash\{\boldsymbol{x}_j\}$ 表示点 $\{\boldsymbol{x}_j\}$ 从设计 Ndesign 中被剔除. CCD 值可以由 (8.19) 的近似表达式计算. 算法 8.4 是一个局部搜索算法, 其可以应用于 R^s 中的任意不规则区域 S.

当不规则区域 S 是单纯形的一部分时, 即

$$S=\left\{\boldsymbol{x}=(x_1,\cdots,x_s)\middle|\sum_{i=1}^{s}x_i=1,x_i\geqslant 0,f_j(\boldsymbol{x})\leqslant 0,j=1,\cdots,t\right\},\tag{8.20}$$

其中 $f_j(\boldsymbol{x})$ 是线性或非线性函数, 此时, Liu Y 和 Liu M Q(2016) 提出另一种构造均匀设计的方法. 实际上, 例 8.6 中的约束实际上是 (8.20) 的特殊情形, 即可以写成线性不等式 $f_j(\boldsymbol{x})\leqslant 0,j=1,\cdots,t$.

Liu Y 和 Liu M Q(2016) 的构造方法是基于单纯形 T^s 的某些变换. 易知, 单纯形 T^s 可以通过等距变换而变换至超平面 $H=\{(z_1,\cdots,z_s)|z_s=0\}$ 中, 即

$$M:\boldsymbol{z}=[\boldsymbol{x}-(1,0,\cdots,0)]Q,\tag{8.21}$$

其中 $\boldsymbol{x} \in T^s$, $\boldsymbol{z} \in H$, Q 是来自于下面的 QR 分解

$$\begin{pmatrix} -1_{1\times(s-1)} \\ I_{s-1} \end{pmatrix} = Q \begin{pmatrix} R_{(s-1)\times(s-1)} \\ 0_{1\times(s-1)} \end{pmatrix}$$

的正交矩阵, 其中 I_{s-1} 是 $s-1$ 阶单位矩阵, $R_{(s-1)\times(s-1)}$ 是上三角矩阵, $-1_{1\times(s-1)}$ 和 $0_{1\times(s-1)}$ 分别为元素都是 -1 和 0 的 $(s-1)$ 维行向量. 单纯形 T^s 中的任意子区域的几何性质都能通过映射 M 在 H 中保留. M 的逆映射可以表示为

$$M^{-1}: \boldsymbol{x} = \boldsymbol{z}Q' + (1, 0, \cdots, 0), \tag{8.22}$$

其中 $\boldsymbol{z} \in M(T^s)$, $\boldsymbol{x} \in T^s$. 基于 CCD 准则, 构造满足约束 (8.20) 的近似均匀设计 (nearly uniform designs, NUD) 的具体算法如下所示 (Liu Y and Liu M Q, 2016).

算法 8.5 (构造单纯形 T^s 子区域上的均匀设计)

步骤 1 设 S 为 (8.20) 定义的单纯形 T^s 的子区域.

步骤 2 通过 (8.21) 的映射 M 把 S 映射到 R^{s-1} 中. 记映射后的区域为 S_1.

步骤 3 给定试验次数 n, 在 CCD 意义下, 通过算法 8.4 的交换算法在 S_1 上得到一个 NUD.

步骤 4 通过 (8.22) 式的逆变换 M^{-1}, 把 NUD 的点映射至单纯形 T^s 中, 从而得到 S 中的 NUD.

算法 8.5 的优点是在于其可以处理一些非常复杂约束的混料设计, 即函数 $f_i(\boldsymbol{x})$ $(i = 1, \cdots, t)$ 可以是非线性的.

例 8.7 (例 8.6 续) 应用算法 8.5 构造不规则区域 S 上的近似均匀设计. 设 $n = 21$. 我们需要得到 (8.21) 式的映射 M 中的正交矩阵 Q. 根据 QR 分解, 有

$$\begin{aligned} \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} &= Q_{3\times 3} \begin{pmatrix} R_{2\times 2} \\ 0_{1\times 2} \end{pmatrix} \\ &= \begin{bmatrix} -\sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{3}/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2}/2 \\ 0 & \sqrt{6}/2 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

则映射 M 和逆映射 M^{-1} 分别为

$$M: \boldsymbol{z} = [\boldsymbol{x} - (1, 0, 0)] \begin{bmatrix} -\sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{3}/3 \end{bmatrix},$$

$$M^{-1}: \boldsymbol{x}=\boldsymbol{z}\begin{bmatrix} -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -\sqrt{6}/6 & -\sqrt{6}/6 & \sqrt{6}/3 \\ \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix}+(1,0,0).$$

根据映射 M, 可以把区域 S 映射到 R^2 中的 S_1, 如图 8.6(b) 所示. 把矩形区域 $[0,\sqrt{2}]\times[0,\sqrt{6}/2]$ 分为 30×30 个同样大小的长方形, 考虑所有长方形的中点并取出落入 S_1 的那些点. 共有 128 个点落入 S_1 中. 令 $N=128$. 给定试验点数 $n=21$, 则由算法 8.4 的交换算法得到的 S_1 中的 NUD 如图 8.7(a) 所示, 由算法 8.5 得到的 S 中相应的 NUD 如图 8.7(b) 所示. 从图 8.7 中可见, 设计点比较均匀地散布在 S 中.

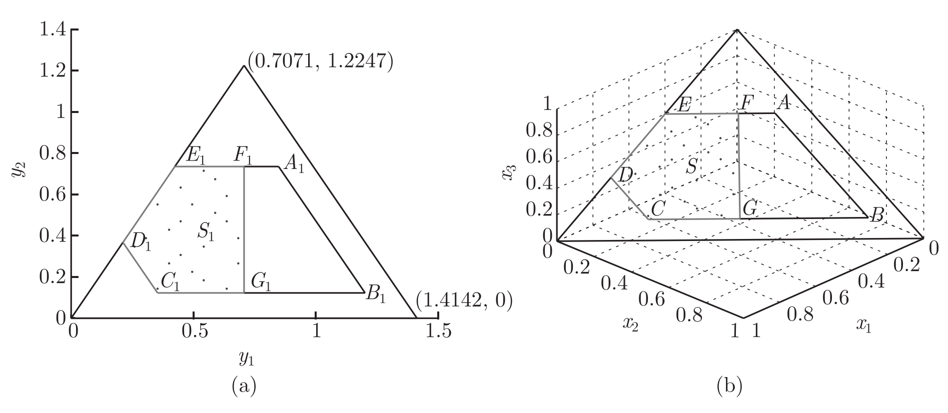


图 8.7　例 8.6 中的均匀设计: (a) S_1 中 $n=21$ 的 NUD; (b) S 中 $n=21$ 的 NUD

8.3　混料设计的建模技术

通常对超立方体上的试验数据, 可以采用一阶模型

$$E(y)=\beta_0+\sum_{i=1}^{s}\beta_i x_i, \tag{8.23}$$

二阶模型

$$E(y)=\beta_0+\sum_{i=1}^{s}\beta_i x_i+\sum_{i=1}^{s}\beta_{ii}x_i^2+\sum_{i<j}^{s}\beta_{ij}x_i x_j \tag{8.24}$$

或中心化的二阶模型

$$E(y)=\beta_0+\sum_{i=1}^{s}\beta_i x_i+\sum_{i=1}^{s}\beta_{ii}(x_i-\bar{x}_i)^2+\sum_{i<j}^{s}\beta_{ij}(x_i-\bar{x}_i)(x_j-\bar{x}_j). \tag{8.25}$$

这些多项式回归模型也常用于混料试验的数据建模. 根据混料试验的约束 $x_1+\cdots+x_s=1$, (8.23) 的一阶模型变为

$$E(y)=\beta_0\sum_{i=1}^{s}x_i+\sum_{i=1}^{s}\beta_i x_i=\sum_{i=1}^{s}\beta_i^* x_i, \tag{8.26}$$

其中 $\beta_i^*=\beta_0+\beta_i, i=1,\cdots,s$. 对于二阶模型 (8.24), 根据 $\sum_{i=1}^{s}x_i=1$ 和 $x_i^2=x_i(1-\sum_{j\neq i}x_j)$, 可得

$$\begin{aligned}E(y)&=\beta_0\sum_{i=1}^{s}x_i+\sum_{i=1}^{s}\beta_i x_i+\sum_{i=1}^{s}\beta_{ii}x_i\left(1-\sum_{j\neq i}x_j\right)+\sum_{i<j}^{s}\beta_{ij}x_i x_j\\&=\sum_{i=1}^{s}(\beta_0+\beta_i+\beta_{ii})x_i-\sum_{i=1}^{s}\beta_{ii}x_i\sum_{j\neq i}x_j+\sum_{i<j}^{s}\beta_{ij}x_i x_j\\&=\sum_{i=1}^{s}\beta_i^* x_i+\sum_{i<j}^{s}\beta_{ij}^* x_i x_j,\end{aligned} \tag{8.27}$$

其中 $\beta_i^*=\beta_0+\beta_i+\beta_{ii}$, $\beta_{ij}^*=\beta_{ij}-\beta_{ii}-\beta_{jj}$. 类似地, 三阶模型也可以重写为

$$E(y)=\sum_{i=1}^{s}\beta_i^* x_i+\sum_{i<j}^{s}\beta_{ij}^* x_i x_j+\sum_{i<j}^{s}\delta_{ij}x_i x_j(x_i-x_j)+\sum_{i<j<k}^{s}\beta_{ijk}^* x_i x_j x_k. \tag{8.28}$$

易知, 模型 (8.26)—(8.28) 中并没有包含截距项、二次项或三次项. 称这些模型为 Scheffé *多项式模型*或*典型多项式形式*. 对于 Scheffé 多项式模型, 可以考虑其最优设计理论. 关于混料设计的最优设计理论可参见 (Chan, 2000).

根据约束 $x_1+\cdots+x_s=1$, 可知 $x_s=1-x_1-\cdots-x_{s-1}$ 且可以把因子 x_s 从模型 (8.23)—(8.25) 中移除, 即下面的模型可以应用于 s 个成分的混料试验

$$E(y)=\beta_0+\sum_{i=1}^{s-1}\beta_i x_i, \tag{8.29}$$

$$E(y)=\beta_0+\sum_{i=1}^{s-1}\beta_i x_i+\sum_{i=1}^{s-1}\beta_{ii}x_i^2+\sum_{i<j}^{s-1}\beta_{ij}x_i x_j, \tag{8.30}$$

$$E(y)=\beta_0+\sum_{i=1}^{s-1}\beta_i x_i+\sum_{i=1}^{s-1}\beta_{ii}(x_i-\bar{x}_i)^2+\sum_{i<j}^{s-1}\beta_{ij}(x_i-\bar{x}_i)(x_j-\bar{x}_j). \tag{8.31}$$

下面的例子将展示对于混料试验的数据建模时可能出现一些新问题. 更多的讨论见文献 (Cornell, 2002).

例 8.8 *在研发新型材料时, 考虑混合三种金属 x_1,x_2,x_3. 采用 15 个试验点的均匀设计. 设计和相应的响应如表 8.5 所示. 由于 $x_1+x_2+x_3=1$, 表中只列出*

x_1 和 x_2. 类似地, 回归模型也只考虑 x_1 和 x_2 这两个因子. 根据变量选择的方法, 考虑下面的模型

$$\begin{aligned}\hat{y} =& 10.472 - 1.20(x_1 - 0.3324) - 3.475(x_1 - 0.3324)^2 \\ & - 3.333(x_2 - 0.3349)^2 + 2.322(x_1 - 0.3324)(x_2 - 0.3349),\end{aligned} \tag{8.32}$$

其中 $R = 0.9887, \hat{\sigma} = 0.07$. 该模型中存在 x_1 和 x_2 之间的交互效应. 图 8.8 给出相应的等高图. 由于因子 x_3 并没有出现在模型 (8.32) 中, 这会给实际应用带来一些不便. (Cornell, 2002) 给出很多讨论以克服这个问题.

表 8.5 设计和响应

编号	x_1	x_2	y	编号	x_1	x_2	y
1	0.81743	0.10346	8.2256	9	0.24723	0.17565	10.1362
2	0.68377	0.05271	8.7794	10	0.20418	0.76930	9.3760
3	0.59175	0.36742	9.5115	11	0.16334	0.25100	10.2772
4	0.51695	0.17712	9.5619	12	0.12440	0.55454	9.8652
5	0.45228	0.41992	9.9145	13	0.08713	0.09129	10.1022
6	0.39447	0.02018	9.5526	14	0.05132	0.79057	9.1792
7	0.34172	0.32914	9.9481	15	0.01681	0.42605	9.9565
8	0.29289	0.49497	10.1241				

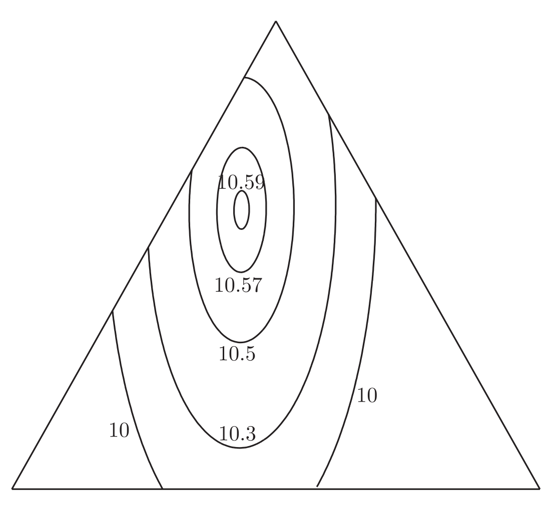


图 8.8 (8.32) 式的等高图

例 8.8 考虑混料设计的建模, 而下面两个问题考虑有约束的混料试验的建模.

例 8.9 (例 8.5 续) 对于表 8.4 的数据, 响应和因子之间的关系是未知的. 我们需要给出一个合适的模型. 从这 12 个响应中可知, 最好的响应值为 $y_6 = 56.6437$

mmHg, 该值可以作为一个基准. 我们需要通过建模来探索是否存在更好的响应.

A. 一阶线性模型

最简单的模型是一阶模型. 基于表 8.4 中的数据, 有

$$\hat{y} = -2.4344 + 79.7565x_1 + 36.4051x_3,$$

其中 $R^2 = 96.87\%$, $s^2 = 4.9638$. 基于方差分析表 (ANOVA)、残差点图、正态图、偏残差点图等统计诊断技术可以说明该一阶模型并不是一个很理想的模型. 因此, 需要考虑更复杂的二阶回归模型.

B. 二阶回归模型

根据模型选择技术, 有下面的回归模型

$$\hat{y}=22.6130 - 11.3570x_2 + 112.205x_1x_3 + 38.0345x_1^2, \tag{8.33}$$

其中 $R^2 = 99.17\%$, $s^2 = 1.4747$. 相应的 ANOVA 结果见表 8.6. 统计诊断中的残差点图和正态图见图 8.9— 图 8.10, 其意味着模型 (8.33) 是可接受的. 在试验区域 $T^3(\boldsymbol{a}, \boldsymbol{b})$ 上选择 $x_i(i = 1, 2, 3)$ 以最大化 y, 可知当 $x_1 = 0.7188, x_2 = 0.0272, x_3 = 0.2540$ 时, 最大值 $\max \hat{y} = 62.4414$.

表 8.6　模型 (8.33) 的 ANOVA 表

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
Model	3	1416.4800	472.1600	320.1736	0.0001
Error	8	11.7976	1.4747		
C Total	11	1428.2776			

Type III Tests					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
X2	1	9.0476	9.0476	6.1352	0.0383
X1 X3	1	75.8957	75.8957	51.4652	0.0001
X1 X1	1	71.9499	71.9499	48.7895	0.0001

C. 中心化二阶回归模型

下面考虑中心化的二阶模型. 再应用模型选择技术, 可得近似模型

$$\begin{aligned}\hat{y}=&37.3126 + 84.6285(x_1 - 0.2775) + 40.9079(x_3 - 0.4565)\\&+84.0939(x_1 - 0.2775)(x_3 - 0.4565),\end{aligned} \tag{8.34}$$

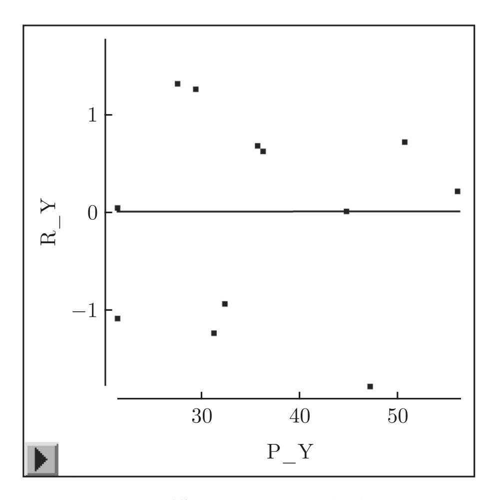


图 8.9 模型 (8.33) 的残差点图

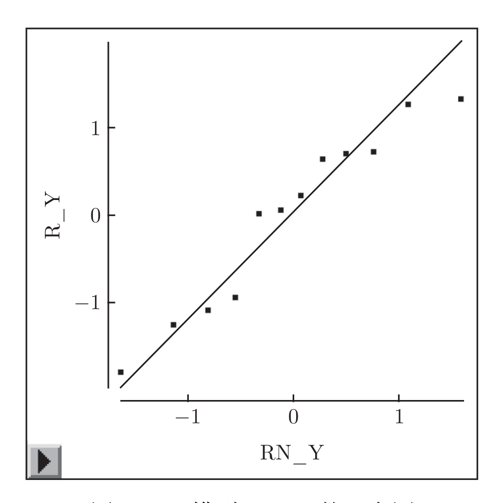


图 8.10 模型 (8.33) 的正态图

其中 $R^2 = 98.94\%$, $s^2 = 1.8969$. 相应的 ANOVA 结果见表 8.7. 残差点图和正态图见图 8.11 和图 8.12. 类似地, 这些图表说明模型 (8.34) 也是可以接受的. 因此, 在模型 (8.34) 下关于 $x_i(i = 1, 2, 3)$ 寻优, 可知当 $x_1 = 0.7188, x_2 = 0.0272, x_3 = 0.2540$ 时, 最优解为 $\max \hat{y} = 59.2179$. 这里最优的成分值与二阶模型是一样的.

在预测的最优点 $x_1 = 0.7188$, $x_2 = 0.0272, x_3 = 0.2540$ 处重复几次试验, 结果发现其平均响应值为 61.75. 从优化和预测的角度来看, 我们认为 (8.33) 是最优的模型, 图 8.13 也给出了模型 (8.33) 的等高图.

表 8.7　模型 (8.34) 的 ANOVA 表

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
Model	3	1413.1024	471.0341	248.3185	0.0001
Error	8	15.1752	1.8969		
C Total	11	1428.2776			

Type III Tests					
Source	DF	Sum of Squares	Mean Square	F Stat	Prob > F
X1	1	1321.3692	1321.3692	696.5958	0.0001
X3	1	307.5655	307.5655	162.1415	0.0001
X1 X3	1	29.4660	29.4660	15.5338	0.0043

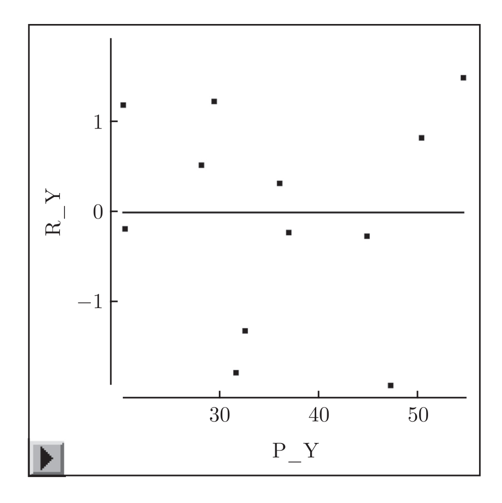


图 8.11　模型 (8.34) 的残差点图

下面的例子来自于 (Tang et al., 2004), 其考虑把混料均匀设计应用于水泥生产.

例 8.10　在建筑行业中经常用到水泥基灌浆材料, 因为其具有高耐久性和高强度. 该材料无毒、无污染并且价格低廉. 然而传统的水泥基灌浆材料也存在一些缺点, 例如, 不稳定、低可加工性、低保水性. 特别地, 当水/水泥的比例较大时, 缺点更凸出. 为克服这些缺点, 有以下几点经验: 适当的添加剂能提高灌浆材料的质量; 硅粉等无机材料能提高混合料的强度、保水性、黏着性, 减少偏析; 羧基甲基纤维素 (carboxyl methyl cellulose, CMC) 等有机材料能提高混合料的黏稠度, 从而减少离析; 粉煤灰将提高混合料的工作性. 粉煤灰是火电厂的一种工业废物, 利用

它可以保护环境. 工程师必须确定每种添加剂的添加量, 这样才能形成一种质量良好、成本效益高的灌浆材料.

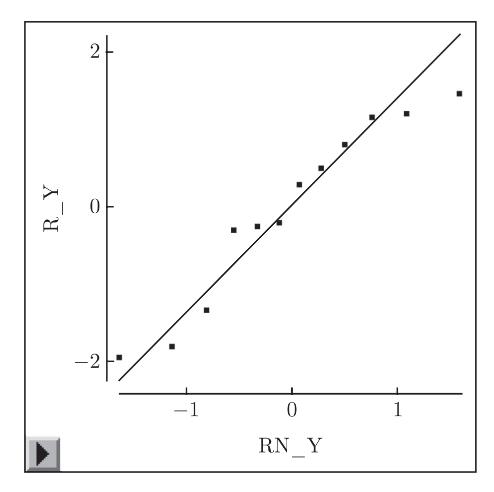


图 8.12 模型 (8.34) 的正态图

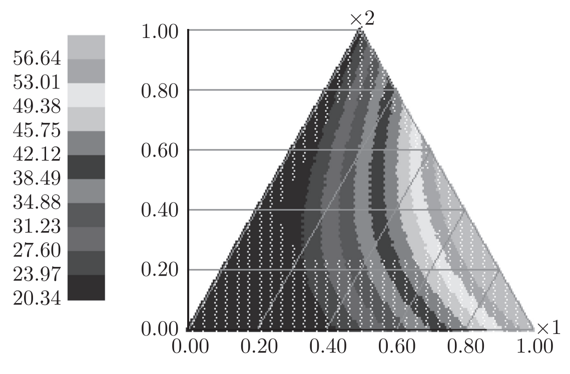


图 8.13 模型 (8.33) 的等高图

我们选择粉煤灰、硅粉、CMC 和水泥作为四个成分, 并分别记为 x_1, x_2, x_3 和 x_4. 试验表明, 前三个成分的百分比的取值范围如下:

$$5 \leqslant x_1 \leqslant 20, \quad 1 \leqslant x_2 \leqslant 2.4, \quad 0.3 \leqslant x_3 \leqslant 1.0.$$

因此, 根据约束 $x_1 + x_2 + x_3 + x_4 = 100$, 水泥 x_4 这一成分的取值范围为 $76.6 \leqslant x_4 \leqslant 93.7$. 这个试验的目的是最小化分散系数 BL (水 / 水泥的比例为 0.6) 和抗压强度 R_{28} (水 / 水泥的比例为 0.8), 后者在水泥混合之后的第 28 天测量.

对于前三个因子, 考虑一个均匀设计 $U_{16}(16 \times 8^2)$, 即 x_1, x_2, x_3 的水平数分别为 16, 8 和 8. 对于前三个成分的某一水平组合, 相应的 x_4 的取值也就确定了. 根

据相应的试验数据, 得到这两个响应与这几个因子之间的关系为

$$\begin{aligned} \text{BL}=&3.337x_1-341.3x_3+0.3655x_4+2.998x_1x_3-0.04575x_1x_4\\ &+5.512x_2x_3-0.0561x_2x_4+3.323x_3x_4, \end{aligned} \tag{8.35}$$

$$\begin{aligned} R_{28}=&-6.694x_1+3.435x_2-4824x_3-0.5854x_1x_2+50.49x_1x_3\\ &+0.0858x_1x_4+61.29x_2x_3+48.53x_3x_4. \end{aligned} \tag{8.36}$$

Tang 等 (2004) 给出关于 BL 和 R_{28} 的一些详细讨论, 并获得最优的材料代价. 因此, 因子设计和正交设计在工业试验设计中得到了广泛的应用. 当因子的数量很大或因子的水平数很大时, 这些设计需要大量的实验次数, 但由于各种约束因子, 在实际中可能无法做到这一点. 在这种情况下, 均匀设计是一个很好的选择, 其可以用于试验.

习　　题

8.1　用 MATLAB 或其他软件编程给出单纯形格子点设计 $\{3,3\},\{4,3\}$ 的试验点及其相应的图形.

8.2　单纯形 T^3 的试验区域实际上是边长为 $\sqrt{2}$ 的等边三角形, 记为 V^2. 因此, V^2 上的任一点 (z_1,z_2) 都对应于 T^3 的某一点 (x_1,x_2,x_3). 选择 V^2 的一个新坐标系, 并给出相应的映射 $(x_1,x_2,x_3)\Rightarrow(z_1,z_2)$.

8.3　对于如下的均匀设计 $U_7(7^2)$:

编号	1	2
1	1	5
2	2	2
3	3	7
4	4	4
5	5	1
6	6	6
7	7	3

应用变换法构造 $T^3=\{(x_1,x_2,x_3):x_i>0,i=1,2,3,x_1+x_2+x_3=1\}$ 上的一个 7 个试验点的均匀设计.

8.4　设 $n=17$.

(1) 随机选取 $[0,1]^2$ 上的 n 个点构造一个设计 D_1, 计算其混合偏差.

(2) 应用变换法得到 T^3 上相应的设计 D_2. 计算 D_2 的均方距离、平均距离和极大距离.

重复步骤 (1) 和 (2) m 次, 比较这些结果并给出你的结论.

8.5　设 $n = 17$. 应用 NTLBG 算法构造 T^3 上的混料均匀设计.

8.6　对于有约束的混料设计, 证明约束 (8.16) 成立.

8.7　考虑例 8.5 中的三个因子. 应用条件法构造一 17 个试验点的有约束的混料均匀设计.

8.8　考虑试验区域

$$S_2 = \{(x_1, x_2, x_3)|x_1 + x_2 + x_3 = 1, x_1^2 + x_2^2 \leqslant 0.36, x_i \geqslant 0\}.$$

在中心复合偏差 CCD 意义下, 应用算法 8.4 中的交换算法构造 S_2 上的一 15 个设计点的均匀设计.

8.9　考虑用混料均匀设计, 来探讨配伍剂量变化对麻杏石甘汤解热作用的影响. 考虑麻黄 (x_1/g)、苦杏仁 (x_2/g)、炙甘草 (x_3/g) 和石膏 (x_4/g) 这四个因子. 在医院里使用的剂量为 (6, 6, 6, 24), 共 42g, 其响应为 6 小时后的热抑制率 (y, %), 该值为 52.19%. 为了研究其他的剂量搭配使得效果更好, 考虑 12 个不同的水平组合的麻杏石甘汤. 每个水平组合的总剂量仍为 42g. 相应的设计点和响应如下所示.

编号	x_1/g	x_2/g	x_3/g	x_4/g	y
1	3.15	25.12	12.02	1.72	41.97
2	17.1	19.82	2.33	2.75	36.13
3	21	2.32	3.89	14.79	28.47
4	1.83	15.57	7.17	17.42	52.92
5	0.59	6.56	18.88	15.97	53.28
6	11.71	16.46	1.73	12.1	29.93
7	14.15	5.83	21.1	0.92	16.79
8	6.09	2.32	26.59	7	29.1
9	7.76	15.75	11.56	6.93	49.64
10	4.56	9.88	1.15	26.41	56.75
11	9.62	0.68	11.89	19.81	52.19
12	27.44	4.7	6.98	2.88	10.53

分析和比较这些数据, 并给出你的结论.

8.10　在某一个中药的试验中, 考虑 5 个成分 x_1—x_5, 其取值范围分别为 10%—60%, 10%—60%, 30%—60%, 10%—12%, 10%—12%. 考虑平均产量 (g) 和存活率 (%) 这两个响应, 分别记为 y_1 和 y_2. 一个 20 个试验点的有约束的混料均匀设计及其响应如下所示.

编号	$x_1/\%$	$x_2/\%$	$x_3/\%$	$x_4/\%$	$x_5/\%$	y_1	y_2
1	15.66	36.69	45.31	0.98	1.36	284.5	44.44
2	33.89	16.77	41.14	1.66	6.53	356.8	44.44
3	19.77	19.03	57.39	1.69	2.11	337.9	88.89
4	36.21	13.36	32.77	8.28	9.37	463.8	100
5	47.08	16.54	33.83	1.28	1.27	326.3	66.66
6	15.57	39.57	35.26	0.65	8.95	454.3	100
7	20.95	33.3	35.23	4.42	6.1	359.1	88.89
8	38.23	14.16	45.3	1.45	0.85	381.4	55.56
9	40.21	16.78	33.03	0.68	9.3	446	55.56
10	17.32	18.97	52.68	0.97	10.05	433.3	77.78
11	18.57	16.08	54.43	9.88	1.04	342.7	66.67
12	31.03	28.74	33.45	5.66	1.12	374	55.56
13	15.96	40.91	32.56	9.72	0.85	397	44.44
14	13.05	34.97	33.66	9.27	9.05	416	88.89
15	14	14.02	52.72	9.33	9.92	475.9	100
16	14.72	33.72	42.64	7.91	1.01	290	22.22
17	32.14	32.36	33.02	0.6	1.88	317.4	88.89
18	40.5	13.49	36.32	9.03	0.66	349.8	44.44
19	26.26	20.98	38.29	8.7	5.77	474.25	44.44
20	15.57	48.33	33.9	1.14	1.06	0	0

分析这些数据并给出最优的成分组合.

参 考 文 献

Amini M M, Alidaee B, Kochenberger G A. 1999. A scatter search approach to unconstrained quadratic binary programs[C]//Corne D, Dorigo M, Glover F. ed. New Ideas in Optimization. McGraw-Hill [s.n.]: 317–329.

Antoniadis A, Oppenheim G. 1995. Wavelets and Statistics[M]. New York: Springer-Verlag.

Aronszajn N. 1950. Theory of reproducing kernels[J]. Trans. Amer. Math. Soc., 68: 337–404.

Atkinson A C, Donev A N. 1992. Optimum Experimental Designs[M]. Oxford: Oxford Science Publications.

Atkinson A C, Bogacka B, Bagacki M B. 1998. D-and T-Optimum designs for the kinetics of a reversible chemical reaction[J]. Chemometr. Intell. Lab. Syst., 43: 185–198.

Bagchi S. 2010. Main-effect plans orthogonal through the block factor[J]. Technometrics, 52: 243–249.

Beasley J E. 1998. Heuristic Algorithms for the Unconstrained Binary Quadratic Programming Problem[R]. London: Management School, Imperial College.

Beth T, Jungnickel D, Lenz H. 1999. Design Theory. Vol. II[M]. Cambridge: Cambridge University Press: i-xxi, 608-1100.

Booth K H V, Cox D. 1962. Some systematic supersaturated designs[J]. Technometrics, 4: 489–495.

Borkowski J J, Piepel G F. 2009. Uniform designs for highly constrained mixture experiments[J]. J. Qual. Technol., 41: 35–37.

Bose M, Bagchi S. 2007. Optimal main effect plans on blocks of small size[J]. Statist. Probab. Lett., 77: 142–147.

Box G. 1963. The effect of error in the factor levels and experimental design[J]. Technometrics, 5: 247–262.

Box G E P, Draper N R. 1959. A basis for the selection of a response surface design[J]. J. Amer. Statist. Assoc., 54: 622–654.

Box G E P, Draper N R. 1987. Empirical Model-Building and Response Surfaces[M]. New York: Wiley.

Box G E P, Hunter J S. 1961a. The 2^{k-p} fractional factorial designs I[J]. Technometrics, 3: 311–351.

Box G E P, Hunter J S. 1961b. The 2^{k-p} fractional factorial designs II[J]. Technometrics, 3: 449–458.

Box G E P, Hunter W G, Hunter J S. 1978. Statistics for Experimenters: An Introduction

to Design, Data Analysis, and Model Building[M]. New York: Wiley.

Boyd S, Vandenberghe L. 2004. Convex Optimization[M]. Cambridge, UK: Cambridge University Press.

Brown L D, Cai T. 1997. Wavelet regression for random uniform design[R]. Department of Statistics, Purdue University.

Bulutoglu D A, Cheng C S. 2004. Construction of $E(s^2)$-optimal supersaturated designs[J]. Ann. Statist., 32: 1662–1678.

Bundschuh P, Zhu Y C. 1993. A Method for Exact Calculation of the Discrepancy of Low-Dimensional Finite Point Sets (I)[C]//Abhandlungen aus Math. Seminar. Bd. 63: Univ.Hamburg.

Butler N A, Mead R, Eskridge K M, et al. 2001. A general method of constructing $E(s^2)$-optimal supersaturated designs[J]. J. R. Stat. Soc. Ser. B, 63: 621–632.

Cai T, Brown L D. 1998. Wavelet shrinkage for nonequispaced samples[J]. Ann. Statist., 26: 1783–1799.

Caliński T, Kageyama S. 2000. Block designs: A randomization Approach. Vol. I[M]. New York: Springer-Verlag.

Cao R Y, Liu M Q. 2015. Construction of second-order orthogonal sliced Latin hypercube designs[J]. J. Complexity, 31: 762–772.

Chan L Y. 2000. Optimal designs for experiments with mixtures: A survey[J]. Comm. Statist. Theory and Methods, 29: 2281–2312.

Chatterjee K, Qin H. 2008. A new look at discrete discrepancy[J]. Statist. Probab. Lett., 78: 2988–2991.

Chatterjee K, Fang K T, Qin H. 2005. Uniformity in factorial designs with mixed levels[J]. J. Statist. Plann. Inference, 128: 593–607.

Chatterjee K, Qin H, Zou N. 2012a. Lee discrepancy on two and three mixed level factorials[J]. Sci. China Ser. A, 55: 663–670.

Chatterjee K, Li Z, Qin H. 2012b. A new lower bound to centered and wrap-round L_2-discrepancies[J]. Statist. Probab. Lett., 82: 1367–1373.

Chen H, Huang H Z, Lin D K J, et al. 2016. Uniform sliced Latin hypercube designs[J]. Appl. Stoch. Models Bus. Ind., 32: 574–584.

Chen J, Lin D K J. 1991. On the identity relationship of 2^{k-p} designs[J]. J. Statist. Plann. Inference, 28: 95–98.

Chen J, Sun D X, Wu C F J. 1993. A catalogue of two-level and three-level fractional factorial designs with small runs[J]. Int. Stat. Rev., 61: 131–145.

Chen J, Liu M Q, Fang K T, et al. 2013. A cyclic construction of saturated and supersaturated designs[J]. J. Statist. Plann. Inference, 143: 2121–2127.

Chen R B, Hsu Y W, Hung Y, et al. 2014. Discrete particle swarm optimization for constructing uniform design on irregular regions[J]. Comput. Statist. Data Anal., 72:

282–297.

Chen V C P, Tsui K L, Barton R R, et al. 2006. A review on design, modeling and applications of computer experiments[J]. IIE Trans., 38: 273–291.

Chen X P, Lin J G, Huang X F. 2015a. Construction of main effects plans orthogonal through the block factor based on level permutation[J]. J. Korean Statist. Soc., 44: 538–545.

Chen X P, Lin J G, Yang J F, et al. 2015b. Construction of main-effect plans orthogonal through the block factor[J]. Statist. Probab. Lett., 106: 58–64.

Chen W, Qi Z F, Zhou Y D. 2015c. Constructing uniform designs under mixture discrepancy[J]. Statist. Probab. Lett., 97: 76–82.

Cheng C S. 1997. $E(s^2)$-optimal supersaturated designs[J]. Statist. Sinica, 7: 929–939.

Cheng C S, Mukerjee R. 1998. Regular fractional factorial designs with minimum aberration and maximum estimation capacity[J]. Ann. Statist., 26: 2289–2300.

Cheng C S, Tang B. 2005. A general theory of minimum aberration and its applications[J]. Ann. Statist., 33: 944–958.

Cheng C S, Deng L W, Tang B. 2002. Generalized minimum aberration and design efficiency for non-regular fractional factorial designs[J]. Statist. Sinica, 12: 991–1000.

Cheng S W, Ye K Q. 2004. Geometric isomorphism and minimum aberration for factorial designs with quantitative factors[J]. Ann. Statist., 32: 2168–2185.

Chuang S C, Hung Y C. 2010. Uniform design over general input domains with applications to target region estimation in computer experiments[J]. Comput. Statist. & Data Anal., 54: 219–232.

Chui C K. 1992. Wavelets: A Tutorial in Theory and Applications[M]. Boston: Academic Press Inc.

Clark J B, Dean A M. 2001. Equivalence of fractional factorial designs[J]. Statist. Sinica, 11: 537–547.

Clerk L D. 1986. A method for exact calculation of the star-discrepancy of plane sets applied to the sequences of Hammersley[J]. Mh. Math., 101: 261–278.

Colbourn C J, Dinita J H. 1996. CRC Handbook of Combinatorial Designs[M]. New York: CRC Press.

Colbourn C J, Dinitz J H, Stinson D R. 1999. Applications of Combinatorial Designs to Communications, Cryptography, and Networking[M]//Surveys in Combinatorics, 1999 (Canterbury). Cambridge: Cambridge University Press. London Math. Soc. Lecture Note Ser., vol. 267.

Cook R D. 1986. Assessment of local influence (with discussion)[J]. J. R. Stat. Soc. Ser. B, 48: 133–169.

Cornell J A. 2002. Experiments with Mixtures, Designs, Models and the Analysis of Mixture Data[M]. 3rd ed. New York: Wiley.

Cornell J A. 2011. A Primer on Experiments with Mixtures[M]. New Jersey, USA: John Wiley & Sons.

Cressie N. 1988. Spatial prediction and ordinary Kriging[J]. Math. Geol., 20: 405–421.

Cressie N. 1993. Statistics for Spatial Data[M]. New York: Wiley.

Daubechies I. 1992. Ten Lectures on Wavelets[M]. Philadelphia: SIAM.

De Boor C, Bon A. 1990. On multivariate polynomial interpolation[J]. Constr. Approx., 6: 287–302.

Dean A, Morris M, Stufken J, et al. 2015. Handbook of Design and Analysis of Experiments[M]. Boca Raton: Chapman and Hall/CRC.

Deng L Y, Tang B. 2002. Design selection and classification for Hadamard matrices using generalized minimum aberration criteria[J]. Technometrics, 44: 173–184.

Dey A. 1986. Theory of Block Designs[M]. New York: Wiley.

Dey A, Mukerjee R. 1999. Fractional Factorial Plans[M]. New York: John Wiley.

Donoho D L, Johnstone I M. 1994. Ideal spatial adaptation by wavelet shrinkage[J]. Biometrika, 81: 425–455.

Donoho D L, Johnstone I M, Kerkyacharian G, et al. 1995. Wavelet shrinkage: Asymptopia?[J]. J. R. Stat. Soc. Ser. B, 57: 301–369.

Draper N R, Beggs W J. 1971. Errors in the factor levels and experimental design[J]. Ann. Math. Statist., 41: 46–58.

Draper N R, Mitchell T J. 1968. Construction of the Set of 256-Run Designs of Resolution $\geqslant 5$ and the Set of Even 512-Run Designs of Resolution $\geqslant 6$ with Special Reference to the Unique Saturated Designs[J]. Ann. Math. Statist., 39: 246–255.

Draper N R, Mitchell T J. 1970. Construction of a Set of 512-Run Designs of Resolution $\geqslant 5$ and a Set of Even 1024-Run Designs of Resolution $\geqslant 6$ [J]. Ann. Math. Statist., 41: 876–887.

Dueck G, Scheuer T. 1990. Threshold accepting: A general purpose algorithm appearing superior to simulated annealing[J]. J. Comput. Phys., 90: 161–175.

Dyn N, Levin D, Rippa S. 1986. Numerical procedures for surface fitting of scattered data by radial basis functions[J]. SIAM J. Sci. and Statist. Computing, 7: 639–659.

Elsawah A M, Qin H. 2014. New Lower Bound for Centered L_2-discrepancy of Four-level U-type Designs[J]. Statist. Probab. Lett., 93: 65–71.

Elsawah A M, Qin H. 2015. Mixture discrepancy on symmetric balanced designs[J]. Statist. Probab. Lett., 104: 123–132.

Elsawah A M, Qin H. 2016. Asymmetric uniform designs based on mixture discrepancy[J]. J. App. Statist., 43(12): 2280–2294.

Elsayed E A. 1996. Reliability Engineering[M]. Massachusetts: Addison Wesley.

Eubank R L. 1988. Spline Smoothing and Nonparametric Regression[M]. New York: Marcel Dekker.

Evangelaras H, Koukouvinos C, Lappas E. 2007. 18-run nonisomorphic three level orthogonal arrays[J]. Metrika, 66: 31–37.

Fan J, Li R. 2001. Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties[J]. J. Amer. Statist. Assoc., 96: 1348–1360.

Fang K T. 1980. The uniform design: Application of number-theoretic methods in experimental design[J]. Acta Math. Appl. Sinica, 3: 363–372.

Fang K T. 2002. Experimental designs for computer experiments and for industrial experiments with model unknown[J]. J. Korean Statist. Soc., 31: 277–299.

Fang K T, Chan L Y. 2006. Uniform design and its industrial applications[C]//Pham H. Handbook of Engineering Statistics. New York: Springer: 229–247.

Fang K T, Ge G N. 2004. An efficient algorithm for the classification of Hadamard matrices[J]. Math. Comp., 73: 843–851.

Fang K T, Hickernell F J. 1995. The Uniform Design and Its Applications[C]//Bulletin of The International Statistical Institute. Beijing: [s.n.]. 50th Session, vol. Book 1.

Fang K T, Li J K. 1995. Some New Results on Uniform Design[J]. Chinese Sci. Bull., 40: 268–272.

Fang K T, Ma C X. 2000. The usefulness of uniformity in experimental design[C]//Kollo T, Tiit E M, Srivastava M. New Trends in Probability and Statistics. The Netherlands: 5.

Fang K T, Ma C X. 2001a. Orthogonal and Uniform Experimental Desigs[M]. Beijing: Science Press.

Fang K T, Ma C X. 2001b. Wrap-Around L_2-Discrepancy of random sampling, Latin hypercube and uniform designs[J]. J. Complexity, 17: 608–624.

Fang K T, Ma C X. 2002. Relationship Between Uniformity, Aberration and Correlation in Regular Fractions 3^{s-1}[C]// Fang K T, Hickernell F J, Niederreiter H. Monte Carlo and Quasi- Monte Carlo Methods 2000.[S.l.]: Springer: 213–231.

Fang K T, Mukerjee R. 2000. A Connection between Uniformity and Aberration in Regular Fractions of Two-level Factorials[J]. Biometrika, 87: 193–198.

Fang K T, Qin H. 2003. A note on construction of nearly uniform designs with large number of runs[J]. Statist. Probab. Lett., 61: 215–224.

Fang K T, Qin H. 2005. Uniformity pattern and related criteria for two-level factorials[J]. Sci. China Ser. A., 48: 1–12.

Fang K T, Wang Y. 1994. Number-Theoretic Methods in Statistics[M]. London: Chapman and Hall.

Fang K T, Winker P. 1998. Uniformity and Orthogonality[R].[S.l.]: Hong Kong Baptist University.

Fang K T, Yang Z H. 2000. On uniform design of experiments with restricted mixtures and generation of uniform distribution on some domains[J]. Statist. Probab. Lett., 46:

113–120.

Fang K T, Zhang A. 2004. Minimum aberration majorization in non-isomorphic saturated designs[J]. J. Statist. Plann. Inference, 126: 337–346.

Fang K T, Kotz S, Ng K W. 1990. Symmetric Multivariate and Related Distributions[M]. London: Chapman and Hall.

Fang K T, Yuan K H, Bentler P M. 1994. Applications of number-theoretic metods to quantizers of elliptically contoured distributions[J]. Multivariate Anal. Appl., 24: 237–251.

Fang K T, Lin D K J, Ma C X. 2000a. On the Construction of Multi-Level Supersaturated Designs[J]. J. Statist. Plann. Inference, 86: 239–252.

Fang K T, Lin D K J, Winker P, et al. 2000b. Uniform Design: Theory and Applications[J]. Technometrics, 42: 237–248.

Fang K T, Ge G N, Liu M Q, et al. 2001. Optimal Supersaturated Designs and Their Constructions[R]. Hong Kong Baptist University.

Fang K T, Ge G N, Liu M Q. 2002a. Construction of $E(f_{\rm NOD})$-Optimal Supersaturated Designs via Room Squares[C]// Chaudhuri A, Ghosh M. Calcutta Statistical Association Bulletin. [S.l.]: 52: 71–84.

Fang K T, Ge G N, Liu M Q. 2002b. Uniform supersaturated design and its construction[J]. Sci. China Ser. A, 45: 1080–1088.

Fang K T, Ma C X, Mukerjee R. 2002c. Uniformity in Fractional Factorials[C]//Fang K T, Hickernell F J, Niederreiter H. Monte Carlo and Quasi-Monte Carlo Methods [S.l.]: New York: Springer, 2000: 232–241.

Fang K T, Ma C X, Winker P. 2002d. Centered L_2-Discrepancy of Random Sampling and Latin Hypercube Design, and Construction of Uniform Designs[J]. Math. Comp., 71: 275–296.

Fang K T, Ge G N, Liu M Q, et al. 2003a. Construction on minimum generalized aberration designs[J]. Metrika, 57: 37–50.

Fang K T, Lu X, Winker P. 2003b. Lower bounds for centered and wrap-around L_2-discrepancies and construction of uniform[J]. J. Complexity, 20: 268–272.

Fang K T, Lin D K J, Liu M Q. 2003c. Optimal mixed-Level supersaturated design[J]. Metrika, 58: 279–291.

Fang K T, Ge G N, Liu M Q, et al. 2004a. Combinatorial constructions for optimal supersaturated designs[J]. Discrete Math., 279: 191–202.

Fang K T, Ge G N, Liu M Q. 2004b. Construction of optimal supersaturated designs by the packing method[J]. Sci. China Ser. A, 47: 128–143.

Fang K T, Ge G N, Liu M Q, et al. 2004c. Construction of uniform designs via super-simple resolvable t-designs[J]. Utilitas Mathematica, 66: 15–32.

Fang K T, Lu X, Tang Y, et al. 2004d. Constructions of uniform designs by using resolvable

packings and coverings[J]. Discrete Math., 274: 25–40.

Fang K T, Tang Y, Yin J X. 2005. Lower bounds for wrap-around L_2-discrepancy and constructions of symmetrical uniform designs[J]. J. Complexity, 21: 757–771.

Fang K T, Li R, Sudjianto A. 2006a. Design and Modeling for Computer Experiments[M]. New York: Chapman and Hall/CRC.

Fang K T, Maringer D, Tang Y, et al. 2006b. Lower bounds and stochastic optimization algorithms for uniform designs with three or four levels[J]. Math. Comp., 75: 859–878.

Fang K T, Tang Y, Yin J X. 2006c. Resolvable partially pairwise balanced designs and their applications in computer experiments[J]. Utilitas Mathematica, 70: 141–157.

Fang K T, Ke X, Elsawah A M. 2016. Construction of a new 27-run uniform orthogonal design[J]. J. Complexity, submitted.

Friedman J H. 1991. Multivariate adaptive regression splines[J]. Ann. Statist., 19: 1–41.

Fries A, Hunter W G. 1980. Minimum aberration 2^{k-p} designs[J]. Technometrics, 22: 601–608.

Georgiou S D. 2014. Supersaturated designs: A review of their construction and analysis[J]. J. Statist. Plann. Inference, 144: 92–109.

Ghosh S, Rao C R. 1996. Handbook of Statistics, 13[M]. New York: North-Holland.

Gilli M, Winkler P. 2009. Heuristic optimization methods in econometrics[C]//Handbook of Computational Econometrics. Belsley, D, Kontoghiorghes E. ed. Wiley, Chichester: 81–119.

Goovaerts P. 1997. Geostatistics for Natural Resources Evaluation[M]. New York: Oxford University Press.

Hall J M. 1961. Hadamard matrix of order 16[J]. Jet Propulsion Laboratory Res. Summery, 1: 21–26.

Hardy R L. 1971. Multiquadric equations of topography and other irregular surfaces[J]. J. Geophys. Res. Atmos., 76: 1905–1915.

Hastie T, Tibshirani R, Friedman J. 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction[M]. New York: Springer.

Hedayat A S, Sloane N J A, Stufken J. 1999. Orthogonal Arrays: Theory and Applications[M]. New York: Springer.

Hickernell F J. 1998a. A generalized discrepancy and quadrature error bound[J]. Math. Comp., 67: 299-322.

Hickernell F J. 1998b. Lattice Rules: How Well Do They Measure Up?[M]//Hellekalek P, Larcher G. Random and Quasi-Random Point Sets.[S.l.]: New York: Springer-Verlag: 106–166.

Hickernell F J, Liu M Q. 2002. Uniform designs limit aliasing[J]. Biometrika, 89: 893–904.

Hiriart-Urruty J B, Lemaréchal C. 1993. Convex Analysis and Minimization Algorithms[M]. New York: Springer.

Ho W M. 2001. Case Studies in Computer Experiments, Applications of Uniform Design and Modern Modeling Techniques[M]. Hong Kong Baptist University: Master Thesis.

Hua L K, Wang Y. 1981. Applications of Number Theory to Numerical Analysis[M]. Berlin and Beijing: Springer and Science Press.

Huang H Z, Yang J F, Liu M Q. 2014. Construction of sliced (nearly) orthogonal Latin hypercube designs[J]. J. Complexity, 30: 355–365.

Huang H Z, Lin D K J, Liu M Q, et al. 2016. Computer experiments with both qualitative and quantitative variables[J]. Technometrics, 58: 495–507.

Huber J. 1975. Robustness and Designs[C]//Srivastave J N. A Survey of Statistical Design and Linear Models. Amsterdam: North-Holland: 287–303.

Iasemidis L D, Pardalos P, Sackellares J C, et al. 2001. Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures[J]. J. Comb. Optim., 5: 9–26.

Jiang B C, Ai M Y. 2014. Construction of uniform designs without replications[J]. J. Complexity, 30: 98–110.

Jin R, Chen W, Sudjianto A. 2005. An Efficient Algorithm for Constructing Optimal Design of Computer Experiments[J]. J. Statist. Plann. Inference, 134: 268–287.

Jing D, Li P, Stagnitti F, et al. 2007. Optimization of laccase production from trametes versicolor by solid fermentation[J]. Can. J. Microbiol., 53: 245–251.

Johnson M E. 1988. Multivariate statistical simulation[J]. J. Roy. Statist. Soc. Ser. A, 151: 930–932.

Johnson M E, Moore L M, Ylvisaker D. 1990. Minimax and maxmin distance designs[J]. J. Statist. Plann. Inference, 26: 131–148.

Katayama K, Narihisa H. 2001. Performance of simulated annealing-based heuristic for the unconstrained binary quadratic programming problem[J]. Eur. J. Operat. Res., 134(1): 103–119.

Ke X, Zhang R, Ye H J. 2015. Two- and three-level lower bounds for mixture L_2-discrepancy and construction of uniform designs by threshold accepting[J]. J. Complexity, 31: 741–753.

Kiefer J. 1959. Optimaml experimental designs (with discussion)[J]. J. Roy. Statist. Soc. Ser. B, 21: 272–319.

Kirkpatrick S, Gelett C, Vecchi M. 1983. Optimization by simulated annealing[J]. Science, 220: 621–630.

Koehler J R, Owen A B. 1996. Computer Experiments[M]//Ghosh S, Rao R C. Handbook of Statistics. Amsterdam: Elsevier Science B. V., 13: 261–308.

Korobov N M. 1959. The approximate computation of multiple integrals[J]. Dokl. Akad. Nauk. SSSR, 124: 1207–1210.

Kozlov M K, Tarasov S P, Khachiyan L G. 1979. Polynomial solvability of convex quadratic

programming[J]. Doklady Akademii Nauk SSSR, 248: 1049–1051, Translated in: Soviet Mathematics–Doklady. 20: 1108–1111.

Li K C, Notz W. 1982. Robust designs for nearly linear regression[J]. J. Statist. Plann. Inference, 6: 135–151.

Li P F, Liu M Q, Zhang R C. 2004. Some theory and the construction of mixed-level supersaturated designs[J]. Statist. Probab. Lett., 69: 105–116.

Li R. 2002. Model selection for analysis of uniform design and computer experiment[J]. Inter. J. Reliability, Quality, and Safety Engineering, 9: 305–315.

Lin D, Draper N R. 1995. Screening properties of certain two-level designs[J]. Metrika, 42: 99–118.

Lin D K J. 1993. A new class of supersaturated designs[J]. Technometrics, 35: 28–31.

Lin D K J, Draper N R. 1992. Projection properties of plackett and burman designs[J]. Technometrics, 34: 423–428.

Linde Y L, Buzo A, et al. 1980. An algorithm for vector quantizer design[J]. IEEE T. Commun., 28: 84–95.

Liu M Q. 2002. Using Discrepancy to Evaluate Fractional Factorial Designs[C]//Fang K T, Hickernell F J, Niederreiter H. Monte Carlo and Quasi-Monte Carlo Methods 2000. Berlin: Springer-Verlag: 357–368.

Liu M Q, Chan L Y. 2004. Uniformity of incomplete block designs[J]. Int. J. Mater. Prod. Tec., 20: 143–149.

Liu M Q, Fang K T. 2005. Some results on resolvable incomplete block designs[J]. Sci. China Ser. A, 48: 503–512.

Liu M Q, Hickernell F J. 2002. $E(s^2)$-Optimality and minimum discrepancy in 2-Level supersaturated designs[J]. Statist. Sinica, 12: 931–939.

Liu M Q, Zhang R C. 2000. Construction of $E(s^2)$ optimal supersaturated designs[J]. J. Statist. Plann. Inference, 86: 229–238.

Liu M Q, Fang K T, Hickernell F J. 2006. Connections among different criteria for asymmetrical fractional factorial designs[J]. Statist. Sinica, 16: 1285–1297.

Liu Y, Dean A. 2004. k-circulant supersaturated designs[J]. Technometrics, 46(1): 32–43.

Liu Y, Liu M Q. 2011. Construction of optimal supersaturated design with large number of levels[J]. J. Statist. Plann. Inference, 141: 2035–2043.

Liu Y, Liu M Q. 2012. Construction of equidistant and weak equidistant supersaturated designs[J]. Metrika, 75: 33–53.

Liu Y, Liu M Q. 2013. Construction of supersaturated design with large number of factors by the complementary design method[J]. Acta Math. Appl. Sinica (English Ser.), 29: 253–262.

Liu Y, Liu M Q. 2016. Construction of uniform designs for mixture experiments with complex constraints[J]. Comm. Statist. Theory Methods, 45: 2172–2180.

Ma C X, Fang K T. 2001. A note on generalized aberration in factorial designs[J]. Metrika, 53: 85–93.

Ma C X, Fang K T. 2004. A new approach to construction of nearly uniform designs[J]. Int. J. Mater. Prod. Tec., 20: 115–126.

Ma C X, Fang K T, Liski E. 2000. A new approach in constructing orthogonal and nearly orthogonal arrays[J]. Metrika, 50: 255–268.

Ma C X, Fang K T, Lin D K J. 2001. On isomorphism of factorial designs[J]. J. Complexity, 17: 86–97.

Ma C X, Fang K T, Lin D K J. 2003. A note on uniformity and orthogonality[J]. J. Statist. Plann. Inference, 113: 323–334.

Marcus M B, Sacks J. 1978. Robust Designs for Regression Problems[C]//Statistical Decision Theory and Related Topics. New York: Academic Press, 2: 245–268.

Marshall A W, Olkin I. 1979. Inequalities: Theory of Majorization and Its Applications[M]. New York: Academic Press.

Matheron G. 1971. The Theory of Regionalized Variables and Its Applications[M]. Mathematiques de Fontainebleau 5. Fontainebleau, France: Les Cahiers du Centre de Morphologie.

McKay M, Beckman R, Conover W. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. Technometrics, 21: 239–245.

Merz P, Freisleben B. 1999. Genetic algorithms for binary quadratic programming[C]// Proceedings of the 1999 Genetic and Evolutionary Computation Conference. [S.l.]: [s.n.]: 417–424.

Merz P, Freisleben B. 2002. Greedy and local search heuristics for unconstrained binary quadratic programming[J]. J. Heuristics, 8: 197–213.

Merz P, Katayama K. 2004. Memetic algorithms for the unconstrained binary quadratic programming problem[J]. Biosystems, 78: 99–118.

Miller A. 2002. Subset Selection in Regression[M]. 2nd. London: Chapman & Hall/CRC.

Miller D, Frenklach M. 1983. Sensitivity analysis and parameter estimation in dynamic modeling of chemical kinetics[J]. Inter. J. Chem. Kinetics, 15: 677–696.

Morris M D. 1991. Factorial sampling plans for preliminary computational experiments[J]. Technometrics, 33: 243–255.

Morris M D, Mitchell T J. 1995. Exploratory design for computational experiments[J]. J. Statist. Plann. Inference, 43: 381–402.

Mukerjee R, Wu C F J. 1995. On the existence of saturated and nearly saturated asymmetrical orthogonal arrays[J]. Ann. Statist., 23: 2102–2115.

Mukerjee R, Wu C F J. 2001. Minimum aberration designs for mixed factorials in terms of complementary sets[J]. Statist. Sinica, 11: 225–239.

Myers R H. 1990. Classical and Modern Regression with Applications[M]. Amsterdam: Duxbury Classic Series.

Mukerjee R, Dey A, Chatterjee K. 2002. Optimal main effect plans with non-orthogonal blocking[J]. Biometrika, 89: 225–229.

Myers R H. 1990. Classical and Modern Regression with Applications[M]. 2nd ed. Belmont, California: Duxbury Press.

Nguyen N K. 1996. An algorithmic approach to constructing supersaturated designs[J]. Technometrics, 38: 69–73.

Niederreiter H. 1973. Application of diophantine approximations to numerical integration[C]//Osgood C F. Diophantine Approximation and Its Applications. New York: Academic Press: 129–199.

Niederreiter H. 1992. Random Number Generation and Quasi-Monte Carlo Methods[C]// SIAM CBMS-NSF Regional Conference. Philadelphia: Applied Mathematics.

Ning J H, Fang K T, Zhou Y D. 2011a. Uniform design for experiments with mixtures[J]. Comm. Statist. Theory Methods, 40: 1734–1742.

Ning J H, Zhou Y D, Fang K T. 2011b. Discrepancy for uniform design of experiments with mixtures[J]. J. Statist. Plann. Inference, 141: 1487–1496.

Owen A B. 1992a. A central limit theorem for latin hypercube sampling[J]. J. R. Statist. Soc. B, 54: 541–551.

Owen A B. 1992b. Randomly orthogonal arrays for computer experiments, integration and visualization[J]. Statist. Sinica, 2: 439–452.

Palubeckis G. 2004. Multistart Tabu search strategies for the unconstrained binary quadratic optimization problem[J]. Ann. Oper. Res., 131: 259–282.

Pardalos P M, Vavasis S A. 1991. Quadratic programming with one negative eigenvalue is NP-hard[J]. J. Global Optim., 1: 15–22.

Patterson H D. 1954. The errors of lattice sampling[J]. J. Roy. Statist. Soc. Ser. B, 16: 140–149.

Piepel G, Anderson C M, Redgate P E. 1993. Response surface designs for irregularly-shaped region - Parts 1, 2, and 3[C]//1993 Proccedings of the Section on Physical and Engineering Sciences. Alexandria, VA: Americal Statistical Associetion: 169–179.

Piepel G, Cooley S, Gan H, et al. 2002. Test matrix support TLCP model development for RPP-WTP HWL glasses[R].[S.l.]: Washington D.C: The Catholic University of America.

Powell M J D. 1987. Radial basis functions for multivariable interpolation: A review[C]// Mason J, Cox M. Algorithms for Approximation. London: Oxford University Press: 143–167.

Prescott P. 2008. Nearly uniform designs for mixture experiments[J]. Comm. Statist. Theory and Methods, 37: 2095–2115.

Pukelsheim F. 1993. Optimum Design of Experiments[M]. New York: Wiley.

Qian P Z G. 2012. Sliced Latin hypercube designs[J]. J. Amer. Statist. Assoc., 107: 393–399.

Qin H, Fang K T. 2004. Discrete discrepancy in factorial designs[J]. Metrika, 60: 59–72H.

Qin H, Li D. 2006. Connection between uniformity and orthogonality for symmetrical factorial designs[J]. J. Stat. Plann. Inference, 136: 2770–2782.

Qin H, Zou N, Zhang S L. 2011. Design efficiency for minimum projection uniform designs with two-level[J]. J. Syst. Sci. Complex., 24: 761–768.

Qin H, Wang Z H, Chatterjee K. 2012. Uniformity pattern and related criteria for q-level factorials[J]. J. Stat. Plann. Inference, 142: 1170–1177.

Riccomango E, Schwabe R, Wynn H P. 1997. Lattice-Based D-Optimal design for Fourier regression[J]. Ann. Statist., 25: 2313–2327.

Roman S. 1992. Coding and Information Theory[M]. New York: Wiley.

Roth R M. 2006. Introduction to Coding Theory[M]. Cambridge, UK: Cambridge University Press.

Sacks J, Schiller S B, Welch W J. 1989a. Designs for computer experiments[J]. Technometrics, 31: 41–47.

Sacks J, Welch W J, Mitchell T J, et al. 1989b. Design and analysis of computer experiments (with discussion)[J]. Statist. Sinica, 4: 409–435.

Santner T J, Williams B J, Notz W I. 2003. The Design and Analysis of Computer Experiments[M]. New York: Springer.

Scheffé H. 1958. Experiments with mixtures[J]. J. R. Stat. Soc. Ser. B., 20: 344–360.

Scheffé H. 1963. The simplex-centroid design for experiments with mixtures[J]. J. R. Stat. Soc. Ser. B., 25: 235–263.

Shaw J E H. 1988. A quasirandom approach to integration in Bayesian statistics[J]. Ann. Statist., 16: 895–914.

Shi P, Fang K T, Tsai C L. 2001. Optimal multi-criteria designs for Fourier regression model[J]. J. Statist. Plann. Inference, 96: 387–401.

Shor N Z. 1985. Minimization Methods for Non-differentiable Functions [M]. Springer Series in Computational Mathematics. New York: Springer.

Shor N Z. 1991. The development of numerical methods for nonsmooth optimization in the USSR[C]//History of Mathematical Programming. A Collection of Personal Reminiscences. Amsterdam: [s.n.] 135–139.

Stein M L. 1987. Large sample properties of simulations using Latin hypercube sampling[J]. Technometrics, 29: 143–151.

Stein M L. 1999. Interpolation of Spatial Data, Some Theory for Kriging[M]. New York: Springer.

Stinson D R. 1991. A survey of Kirkman triple systems and related designs[J]. Discrete

Math., 92(1-3): 371–393.

Stinson D R. 1996. Packings[C]//Colbourn C J, Dinitz J H. CRC Press Series on Discrete Math. and its Applications. Boca Raton, FL: CRC Press: 409–413.

Sun D X, Wu C F J. 1993. Statistical properties of Hadamard matrices of order 16[C]//Kuo W. Quality through Engineering Design. Amsterdam: Elsevier: 169–179.

Sun F S, Liu M Q, Hao W R. 2009. An algorithmic approach to finding factorial designs with generalized minimum aberration[J]. J. Complexity, 25: 75–84.

Sun F S, Chen J, Liu M Q. 2011a. Connections between uniformity and aberration in general multi-level factorials[J]. Metrika, 73: 305-315.

Sun F S, Lin D K J, Liu M Q. 2011b. On construction of optimal mixed-level supersaturated designs[J]. Ann. Statist., 39: 1310–1333.

Tang B. 1993. Orthogonal array-Based Latin hypercubes[J]. J. Amer. Statat. Assoc., 88: 1392–1397.

Tang B. 2001. Theory of J-characteristics for fractional factorial designs and projection justification of minimum G_2-aberration[J]. Biometrika, 88: 401–407.

Tang B, Deng L Y. 1999. Minimum G_2-Aberration for nonregular fractional designs[J]. Ann. Statist., 27: 1914–1926.

Tang B, Wu C F J. 1997. A method for constructing supersaturated designs and its $E(s^2)$ optimality[J]. Canad. J. Statist., 25: 191–201.

Tang M, Li J, Chan L Y, et al. 2004. Application of uniform design in the formation of cement mixtures[J]. Qua. Eng., 16: 461–474.

Tang Y. 2005. Combinatorial Properties of Uniform Designs and Their Applications in the Constructions of Low-discrepancy Designs[R]. Hong Kong Baptist University.

Tang Y, Xu H. 2013. An effective construction method for multilevel uniform designs[J]. J. Statist. Plann. Inference, 143: 1583–1589.

Tang Y, Xu H, Lin D K J. 2012. Uniform fractional factorial designs[J]. Ann. Statist., 40: 891–907.

Wahba G. 1990. Spline Models for Observational Data[M]. Philadelphia, PA, USA: SIAM.

Wang J C, Wu C F J. 1992. Nearly orthogonal arrays with mixed levels and small runs[J]. Technometrics, 34: 409–422.

Wang Y, Fang K T. 1990. Number theoretic methods in applied statistics (II)[J]. Chinese Ann. Math. Ser. B., 11: 384–395.

Wang Y, Fang K T. 1996. Uniform design of experiments with mixtures[J]. Sci. China Ser. A, 39: 264–275.

Warnock T T. 1972. Computational Investigations of Low Discrepancy Point Sets[C]// Zaremba S K. Applications of Number Theory to Numerical Analysis. New York: Academic Press: 319–343.

Welch W J, Buck R J, Sacks J, et al. 1992. Screening, predicting and computer experi-

ments[J]. Technometrics, 34: 15–25.

Weyl H. 1916. Über die Gleichverteilung der Zahlem mod Eins[J]. Math. Ann., 77: 313–352.

Wiens D P. 1990. Robust minimax designs for multiple linear regression[J]. Linear Algebra Appl., 127: 327–340.

Wiens D P. 1991. Designs for approximately linear regression: Two optimality properties of uniform designs[J]. Statist. Probab. Lett., 12: 217–221.

Wiens D P. 1992. Minimax designs for approximately linear regression[J]. J. Statist. Plann. Inference, 31: 353–371.

Winker P, Fang K T. 1997. Application of threshold accepting to the evaluation of the discrepancy of a set of points[J]. SIAM Numer. Anal., 34: 2038–2042.

Winker P, Fang K T. 1998. Optimal U-type design[C]//Niederreiter H, Zinterhof P, Hellekalek P. Monte Carlo and Quasi-Monte Carlo Methods 1996.[S.l.]: Springer: 436–448.

Wu C F J, Hamada M. 2009. Experiments: Planning, Analysis, and Optimization[M]. 2nd ed. New York: John Wiley and Sons.

Xie M Y, Fang K T. 2000. Admissibility and minimaxity of the uniform design measure in nonparametric regression model[J]. J. Statist. Plann. Inference, 83: 101–111.

Xie M Y, Ning J H, Fang K T. 2007. Orthogonality and D-optimality of the U-type design under general Fourier regression models[J]. Statist. Probab. Lett., 77: 1377–1384.

Xu H. 2003. Minimum moment aberration for nonregular designs and supersaturated designs[J]. Statist. Sinica, 13: 691–708.

Xu H, Wu C F J. 2001. Generalized minimum aberration for asymmetrical fractional factorial designs[J]. Ann. Statist., 29: 549–560.

Xu Q S, Liang Y Z, Fang K T. 2000. The effects of different experimental designs on parameter estimation in the kinetics of a reversible chemical reaction[J]. Chemometrics and Intelligent Laboratory System, 52: 155–166.

Yamada S, Lin D K J. 1999. 3-Level supersaturated designs[J]. Statist. Probab. Lett., 45: 31–39.

Yamada S, Matsui T. 2002. Optimality of mixed-level supersaturated designs[J]. J. Statist. Plann. Inference, 104(2): 459–468.

Yamada S, Ikebe Y T, Hashiguchi H, et al. 1999. Construction of three-Level supersaturated design[J]. J. Statist. Plann. Inference, 81: 183–193.

Yang J F, Sun F S, Lin D K J, et al. 2010. A study on design uniformity under errors in the level values[J]. Statist. Probab. Lett., 80: 1467–1471.

Yang J F, Lin C D, Qian P Z G, et al. 2013. Construction of sliced orthogonal Latin hypercube designs[J]. Statist. Sinica, 23: 1117–1130.

Yang J Y, Chen H, Lin D K J, et al. 2016. Construction of sliced maximin-orthogonal

Latin hypercube designs[J]. Statist. Sinica, 26: 589–603.

Yang X, Chen H, Liu M Q. 2014. Resolvable orthogonal array-based uniform sliced Latin hypercube designs[J]. Statist. Probab. Lett., 93: 108–115.

Ye K Q. 2003. Indicator fundction and its appliction in two-level factorial design[J]. Ann. Statist., 31: 984–994.

Yin Y H, Lin D K J, Liu M Q. 2014. Sliced Latin hypercube designs via orthogonal arrays[J]. J. Statist. Plann. Inference, 149: 162–171.

Yuan R, Lin D K J, Liu M Q. 2017. Nearly column-orthogonal designs based on leave-one-out good lattice point sets[J]. J. Statist. Plann. Inference, 185: 29–40.

Yue R X, Hickernell F J. 1999. Robust designs for fitting liner models with misspecification[J]. Statist. Sinica, 9.

Zeng Z J. 1994. The Uniform Design and Its Applications[M]. (in Chinese) Shenyang: Liaoning People's Publishing House.

Zhang A, Fang K T, Li R, et al. 2005. Majorization Framework for Balanced Lattice Designs[J]. Ann. Statist., 33: 2837–2853.

Zhang Q, Wang Z, Hu J, et al. 2015. A new lower bound for wrap-around L_2-discrepancy on two and three mixed level factorials[J]. Statist. Probab. Lett., 96: 133–140.

Zhang S L, Qin H. 2006. Minimum projection uniformity criterion and its application[J]. Statist. Probab. Lett., 76: 634–640.

Zhou Y D, Fang K T. 2013. An efficient method for constructing uniform designs with large size[J]. Comput. Statist., 28(3): 1319–1331.

Zhou Y D, Ning J H. 2008. Lower bounds of the wrap-around L_2-discrepancy and relationships between Mlhd and uniform design with a large size[J]. J. Stat. Plann. Inference, 138: 2330–2339.

Zhou Y D, Xu H. 2014. Space-Filling fractional factorial designs[J]. J. Amer. Statist. Assoc., 109: 1134–1144.

Zhou Y D, Xu H. 2015. Space-filling properties of good lattice point sets[J]. Biometrika, 102: 959–966.

Zhou Y D, Ning J H, Song X B. 2008. Lee discrepancy and its applications in experimental designs[J]. Statist. Probab. Lett., 78: 1933–1942.

Zhou Y D, Fang K T, Ning J H. 2012. Constructing uniform designs: A heuristic integer programming method[J]. J. Complexity, 28: 224–237.

Zhou Y D, Fang K T, Ning J H. 2013. Mixture discrepancy for quasi-random point sets[J]. J. Complexity, 29: 283–301.

Zou N, Ren P, Qin H. 2009. A note on Lee discrepancy[J]. Statist. Probab. Lett., 79: 496–500.

索　引

《现代数学基础丛书》已出版书目

(按出版时间排序)

1　数理逻辑基础(上册)　1981.1　胡世华　陆钟万　著
2　紧黎曼曲面引论　1981.3　伍鸿熙　吕以辇　陈志华　著
3　组合论(上册)　1981.10　柯　召　魏万迪　著
4　数理统计引论　1981.11　陈希孺　著
5　多元统计分析引论　1982.6　张尧庭　方开泰　著
6　概率论基础　1982.8　严士健　王隽骧　刘秀芳　著
7　数理逻辑基础(下册)　1982. 8　胡世华　陆钟万　著
8　有限群构造(上册)　1982.11　张远达　著
9　有限群构造(下册)　1982.12　张远达　著
10　环与代数　1983.3　刘绍学　著
11　测度论基础　1983.9　朱成熹　著
12　分析概率论　1984.4　胡迪鹤　著
13　巴拿赫空间引论　1984.8　定光桂　著
14　微分方程定性理论　1985.5　张芷芬　丁同仁　黄文灶　董镇喜　著
15　傅里叶积分算子理论及其应用　1985.9　仇庆久等　编
16　辛几何引论　1986.3　J.柯歇尔　邹异明　著
17　概率论基础和随机过程　1986.6　王寿仁　著
18　算子代数　1986.6　李炳仁　著
19　线性偏微分算子引论(上册)　1986.8　齐民友　著
20　实用微分几何引论　1986.11　苏步青等　著
21　微分动力系统原理　1987.2　张筑生　著
22　线性代数群表示导论(上册)　1987.2　曹锡华等　著
23　模型论基础　1987.8　王世强　著
24　递归论　1987.11　莫绍揆　著
25　有限群导引(上册)　1987.12　徐明曜　著
26　组合论(下册)　1987.12　柯　召　魏万迪　著
27　拟共形映射及其在黎曼曲面论中的应用　1988.1　李　忠　著
28　代数体函数与常微分方程　1988.2　何育赞　著
29　同调代数　1988.2　周伯壎　著

30　近代调和分析方法及其应用　1988.6　韩永生　著
31　带有时滞的动力系统的稳定性　1989.10　秦元勋等　编著
32　代数拓扑与示性类　1989.11　马德森著　吴英青　段海鲍译
33　非线性发展方程　1989.12　李大潜　陈韵梅　著
34　反应扩散方程引论　1990.2　叶其孝等　著
35　仿微分算子引论　1990.2　陈恕行等　编
36　公理集合论导引　1991.1　张锦文　著
37　解析数论基础　1991.2　潘承洞等　著
38　拓扑群引论　1991.3　黎景辉　冯绪宁　著
39　二阶椭圆型方程与椭圆型方程组　1991.4　陈亚浙　吴兰成　著
40　黎曼曲面　1991.4　吕以辇　张学莲　著
41　线性偏微分算子引论(下册)　1992.1　齐民友　许超江　编著
42　复变函数逼近论　1992.3　沈燮昌　著
43　Banach 代数　1992.11　李炳仁　著
44　随机点过程及其应用　1992.12　邓永录等　著
45　丢番图逼近引论　1993.4　朱尧辰等　著
46　线性微分方程的非线性扰动　1994.2　徐登洲　马如云　著
47　广义哈密顿系统理论及其应用　1994.12　李继彬　赵晓华　刘正荣　著
48　线性整数规划的数学基础　1995.2　马仲蕃　著
49　单复变函数论中的几个论题　1995.8　庄圻泰　著
50　复解析动力系统　1995.10　吕以辇　著
51　组合矩阵论　1996.3　柳柏濂　著
52　Banach 空间中的非线性逼近理论　1997.5　徐士英　李　冲　杨文善　著
53　有限典型群子空间轨道生成的格　1997.6　万哲先　霍元极　著
54　实分析导论　1998.2　丁传松等　著
55　对称性分岔理论基础　1998.3　唐　云　著
56　Gel'fond-Baker 方法在丢番图方程中的应用　1998.10　乐茂华　著
57　半群的 S-系理论　1999.2　刘仲奎　著
58　有限群导引(下册)　1999.5　徐明曜等　著
59　随机模型的密度演化方法　1999.6　史定华　著
60　非线性偏微分复方程　1999.6　闻国椿　著
61　复合算子理论　1999.8　徐宪民　著
62　离散鞅及其应用　1999.9　史及民　编著
63　调和分析及其在偏微分方程中的应用　1999.10　苗长兴　著

64　惯性流形与近似惯性流形　2000.1　戴正德　郭柏灵　著
65　数学规划导论　2000.6　徐增堃　著
66　拓扑空间中的反例　2000.6　汪　林　杨富春　编著
67　拓扑空间论　2000.7　高国士　著
68　非经典数理逻辑与近似推理　2000.9　王国俊　著
69　序半群引论　2001.1　谢祥云　著
70　动力系统的定性与分支理论　2001.2　罗定军　张　祥　董梅芳　编著
71　随机分析学基础(第二版)　2001.3　黄志远　著
72　非线性动力系统分析引论　2001.9　盛昭瀚　马军海　著
73　高斯过程的样本轨道性质　2001.11　林正炎　陆传荣　张立新　著
74　数组合地图论　2001.11　刘彦佩　著
75　光滑映射的奇点理论　2002.1　李养成　著
76　动力系统的周期解与分支理论　2002.4　韩茂安　著
77　神经动力学模型方法和应用　2002.4　阮炯　顾凡及　蔡志杰　编著
78　同调论 —— 代数拓扑之一　2002.7　沈信耀　著
79　金兹堡-朗道方程　2002.8　郭柏灵等　著
80　排队论基础　2002.10　孙荣恒　李建平　著
81　算子代数上线性映射引论　2002.12　侯晋川　崔建莲　著
82　微分方法中的变分方法　2003.2　陆文端　著
83　周期小波及其应用　2003.3　彭思龙　李登峰　谌秋辉　著
84　集值分析　2003.8　李　雷　吴从炘　著
85　数理逻辑引论与归结原理　2003.8　王国俊　著
86　强偏差定理与分析方法　2003.8　刘　文　著
87　椭圆与抛物型方程引论　2003.9　伍卓群　尹景学　王春朋　著
88　有限典型群子空间轨道生成的格(第二版)　2003.10　万哲先　霍元极　著
89　调和分析及其在偏微分方程中的应用(第二版)　2004.3　苗长兴　著
90　稳定性和单纯性理论　2004.6　史念东　著
91　发展方程数值计算方法　2004.6　黄明游　编著
92　传染病动力学的数学建模与研究　2004.8　马知恩　周义仓　王稳地　靳　祯　著
93　模李超代数　2004.9　张永正　刘文德　著
94　巴拿赫空间中算子广义逆理论及其应用　2005.1　王玉文　著
95　巴拿赫空间结构和算子理想　2005.3　钟怀杰　著
96　脉冲微分系统引论　2005.3　傅希林　闫宝强　刘衍胜　著
97　代数学中的 Frobenius 结构　2005.7　汪明义　著

98　生存数据统计分析　2005.12　王启华　著
99　数理逻辑引论与归结原理(第二版)　2006.3　王国俊　著
100　数据包络分析　2006.3　魏权龄　著
101　代数群引论　2006.9　黎景辉　陈志杰　赵春来　著
102　矩阵结合方案　2006.9　王仰贤　霍元极　麻常利　著
103　椭圆曲线公钥密码导引　2006.10　祝跃飞　张亚娟　著
104　椭圆与超椭圆曲线公钥密码的理论与实现　2006.12　王学理　裴定一　著
105　散乱数据拟合的模型方法和理论　2007.1　吴宗敏　著
106　非线性演化方程的稳定性与分歧　2007.4　马　天　汪守宏　著
107　正规族理论及其应用　2007.4　顾永兴　庞学诚　方明亮　著
108　组合网络理论　2007.5　徐俊明　著
109　矩阵的半张量积:理论与应用　2007.5　程代展　齐洪胜　著
110　鞅与 Banach 空间几何学　2007.5　刘培德　著
111　非线性常微分方程边值问题　2007.6　葛渭高　著
112　戴维-斯特瓦尔松方程　2007.5　戴正德　蒋慕蓉　李栋龙　著
113　广义哈密顿系统理论及其应用　2007.7　李继彬　赵晓华　刘正荣　著
114　Adams 谱序列和球面稳定同伦群　2007.7　林金坤　著
115　矩阵理论及其应用　2007.8　陈公宁　著
116　集值随机过程引论　2007.8　张文修　李寿梅　汪振鹏　高　勇　著
117　偏微分方程的调和分析方法　2008.1　苗长兴　张　波　著
118　拓扑动力系统概论　2008.1　叶向东　黄　文　邵　松　著
119　线性微分方程的非线性扰动(第二版)　2008.3　徐登洲　马如云　著
120　数组合地图论(第二版)　2008.3　刘彦佩　著
121　半群的 S-系理论(第二版)　2008.3　刘仲奎　乔虎生　著
122　巴拿赫空间引论(第二版)　2008.4　定光桂　著
123　拓扑空间论(第二版)　2008.4　高国士　著
124　非经典数理逻辑与近似推理(第二版)　2008.5　王国俊　著
125　非参数蒙特卡罗检验及其应用　2008.8　朱力行　许王莉　著
126　Camassa-Holm 方程　2008.8　郭柏灵　田立新　杨灵娥　殷朝阳　著
127　环与代数(第二版)　2009.1　刘绍学　郭晋云　朱　彬　韩　阳　著
128　泛函微分方程的相空间理论及应用　2009.4　王　克　范　猛　著
129　概率论基础(第二版)　2009.8　严士健　王隽骧　刘秀芳　著
130　自相似集的结构　2010.1　周作领　瞿成勤　朱智伟　著
131　现代统计研究基础　2010.3　王启华　史宁中　耿　直　主编

132　图的可嵌入性理论(第二版)　2010.3　刘彦佩　著
133　非线性波动方程的现代方法(第二版)　2010.4　苗长兴　著
134　算子代数与非交换 L_p 空间引论　2010.5　许全华　吐尔德别克　陈泽乾　著
135　非线性椭圆型方程　2010.7　王明新　著
136　流形拓扑学　2010.8　马　天　著
137　局部域上的调和分析与分形分析及其应用　2011.6　苏维宜　著
138　Zakharov 方程及其孤立波解　2011.6　郭柏灵　甘在会　张景军　著
139　反应扩散方程引论(第二版)　2011.9　叶其孝　李正元　王明新　吴雅萍　著
140　代数模型论引论　2011.10　史念东　著
141　拓扑动力系统——从拓扑方法到遍历理论方法　2011.12　周作领　尹建东　许绍元　著
142　Littlewood-Paley 理论及其在流体动力学方程中的应用　2012.3　苗长兴　吴家宏　章志飞　著
143　有约束条件的统计推断及其应用　2012.3　王金德　著
144　混沌、Mel′nikov 方法及新发展　2012.6　李继彬　陈凤娟　著
145　现代统计模型　2012.6　薛留根　著
146　金融数学引论　2012.7　严加安　著
147　零过多数据的统计分析及其应用　2013.1　解锋昌　韦博成　林金官　编著
148　分形分析引论　2013.6　胡家信　著
149　索伯列夫空间导论　2013.8　陈国旺　编著
150　广义估计方程估计方法　2013.8　周　勇　著
151　统计质量控制图理论与方法　2013.8　王兆军　邹长亮　李忠华　著
152　有限群初步　2014.1　徐明曜　著
153　拓扑群引论(第二版)　2014.3　黎景辉　冯绪宁　著
154　现代非参数统计　2015.1　薛留根　著
155　三角范畴与导出范畴　2015.5　章　璞　著
156　线性算子的谱分析(第二版)　2015.6　孙　炯　王　忠　王万义　编著
157　双周期弹性断裂理论　2015.6　李　星　路见可　著
158　电磁流体动力学方程与奇异摄动理论　2015.8　王　术　冯跃红　著
159　算法数论(第二版)　2015.9　裴定一　祝跃飞　编著
160　偏微分方程现代理论引论　2016.1　崔尚斌　著
161　有限集上的映射与动态过程——矩阵半张量积方法　2015.11　程代展　齐洪胜　贺风华　著
162　现代测量误差模型　2016.3　李高荣　张　君　冯三营　著
163　偏微分方程引论　2016. 3　韩丕功　刘朝霞　著
164　半导体偏微分方程引论　2016.4　张凯军　胡海丰　著
165　散乱数据拟合的模型、方法和理论(第二版)　2016.6　吴宗敏　著

166 交换代数与同调代数(第二版) 2016.12 李克正 著

167 Lipschitz 边界上的奇异积分与 Fourier 理论 2017.3 钱 涛 李澎涛 著

168 有限 p 群构造(上册) 2017.5 张勤海 安立坚 著

169 有限 p 群构造(下册) 2017.5 张勤海 安立坚 著

170 自然边界积分方法及其应用 2017.6 余德浩 著

171 非线性高阶发展方程 2017.6 陈国旺 陈翔英 著

172 数理逻辑导引 2017.9 冯 琦 编著

173 简明李群 2017.12 孟道骥 史毅茜 著

174 代数 K 理论 2018.6 黎景辉 著

175 线性代数导引 2018.9 冯 琦 编著

176 基于框架理论的图像融合 2019.6 杨小远 石 岩 王敬凯 著

177 均匀试验设计的理论和应用 2019.10 方开泰 刘民千 覃 红 周永道 著

彩　　图

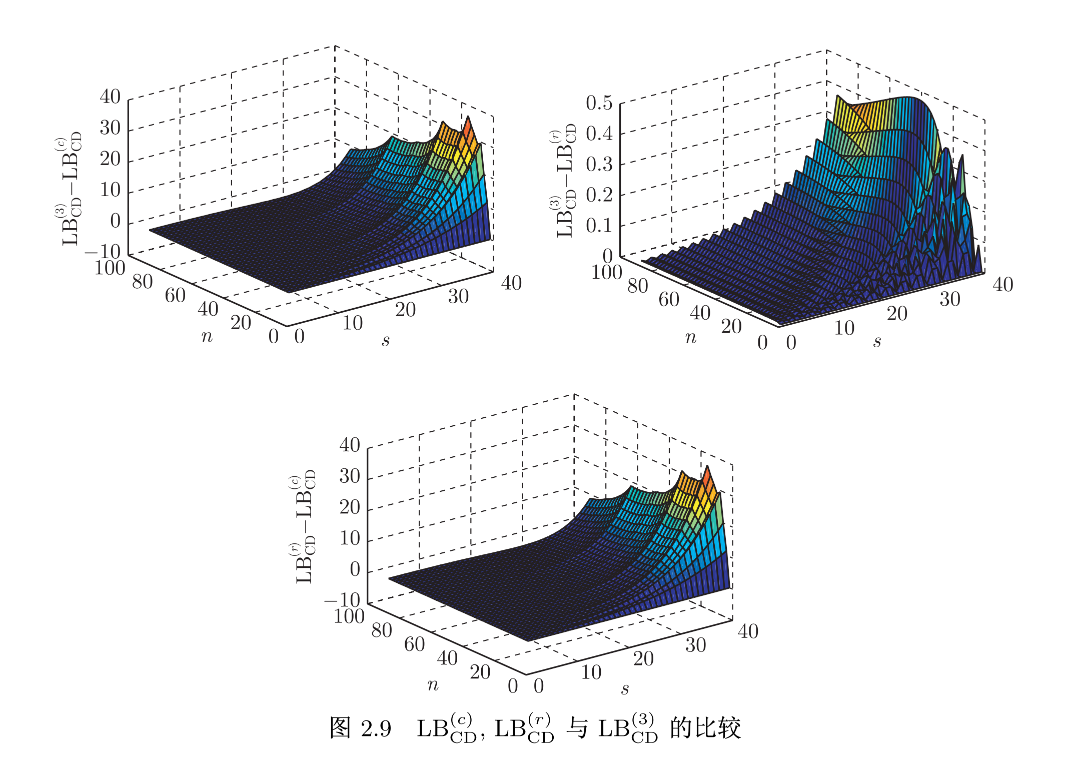


图 2.9　$\mathrm{LB}_{\mathrm{CD}}^{(c)}$, $\mathrm{LB}_{\mathrm{CD}}^{(r)}$ 与 $\mathrm{LB}_{\mathrm{CD}}^{(3)}$ 的比较

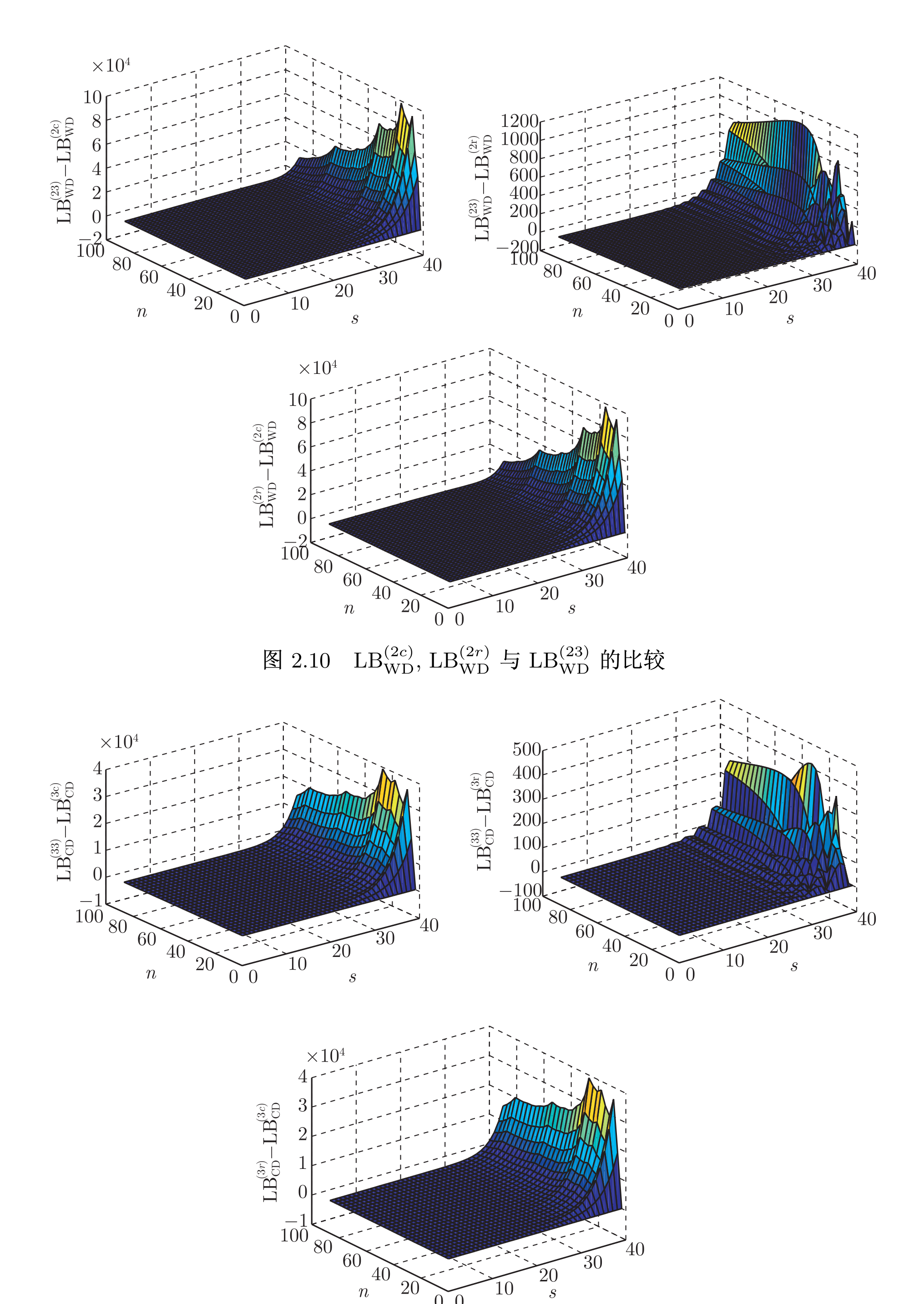
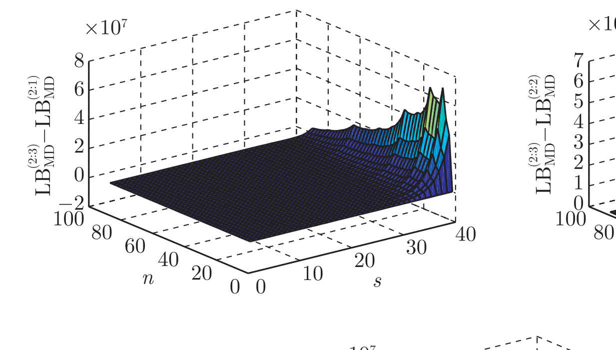
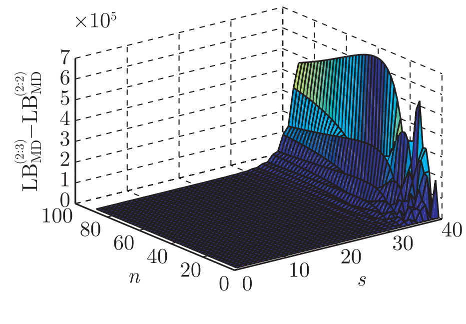


图 2.10　$\mathrm{LB}_{\mathrm{WD}}^{(2c)}, \mathrm{LB}_{\mathrm{WD}}^{(2r)}$ 与 $\mathrm{LB}_{\mathrm{WD}}^{(23)}$ 的比较

图 2.11　$\mathrm{LB}_{\mathrm{WD}}^{(3c)}, \mathrm{LB}_{\mathrm{WD}}^{(3r)}$ 与 $\mathrm{LB}_{\mathrm{WD}}^{(33)}$ 的比较





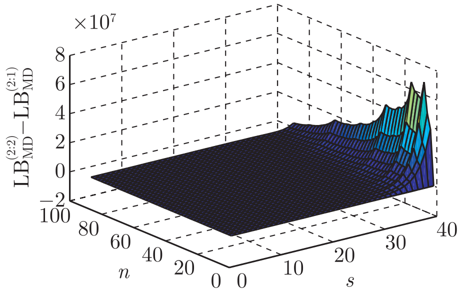


图 2.12　$LB_{MD}^{(2:1)}$, $LB_{MD}^{(2:2)}$ 与 $LB_{MD}^{(2:3)}$ 的比较

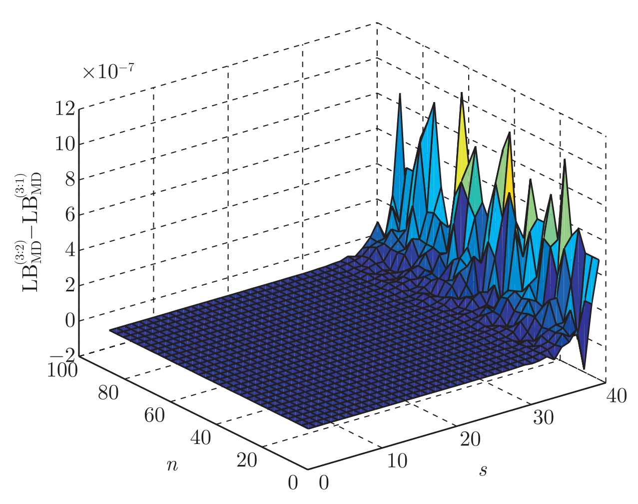


图 2.13　$LB_{MD}^{(3:1)}$ 与 $LB_{MD}^{(3:2)}$ 的比较

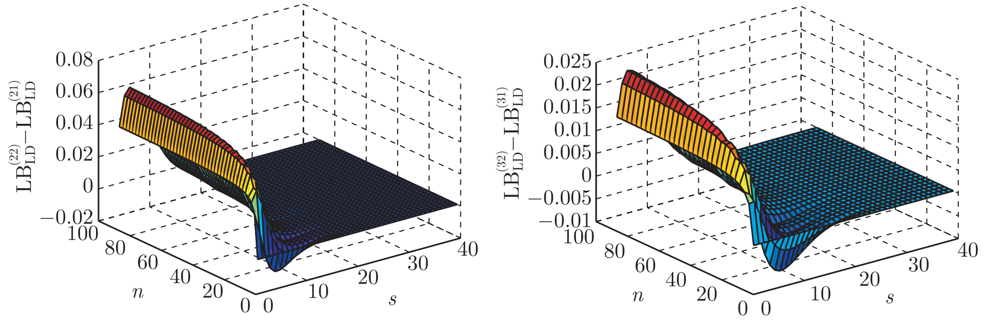


图 2.14　$LB_{LD}^{(21)}$ 与 $LB_{LD}^{(22)}$, $LB_{LD}^{(31)}$ 与 $LB_{LD}^{(32)}$ 的比较

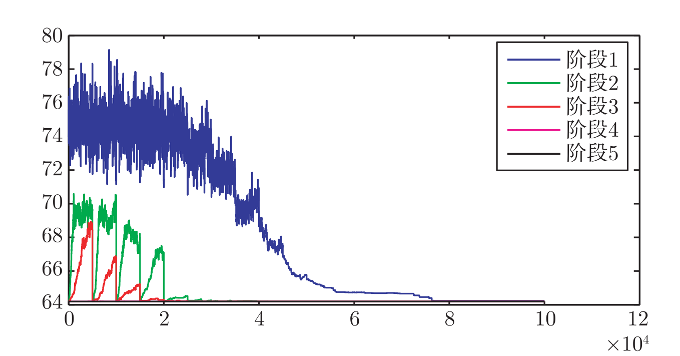


图 4.4　例 4.1 中多阶段 TA 算法在不同阶段的混合偏差的轨迹图

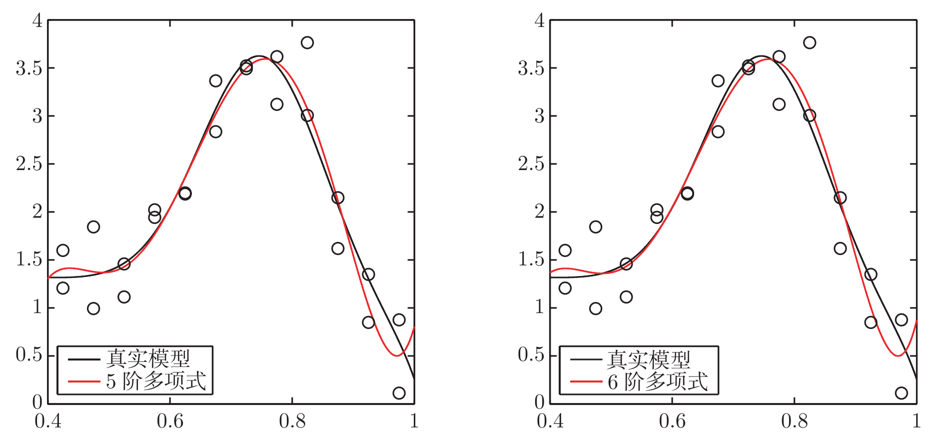


图 5.1　5 阶和 6 阶中心化多项式回归

(O-7821.31)



科学出版社互联网入口
科学数理分社
电 话：(010) 64019814
Email：lijingke@mail.sciencep.com
销售分类建议：高等数学

www.sciencep.com



定 价：128.00 元